(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 972 830 A1

(12) EUROPEAN PATENT APPLICATION

published in accordance with Art. 158(3) EPC

(43) Date of publication:
19.01.2000 Bulletin 2000/03

(21) Application number: **98904368.2**

(22) Date of filing: **18.02.1998**

(51) Int. Cl.[7]: **C12N 15/11**, C12Q 1/68, C07K 14/47, C07K 16/18 // (C12N15/11, C12R1:91)

(86) International application number:
PCT/JP98/00667

(87) International publication number:
WO 98/37187 (27.08.1998 Gazette 1998/34)

(84) Designated Contracting States:
DE FR GB IT NL

(30) Priority: **21.02.1997 JP 5250897**

(71) Applicant: **Takara Shuzo Co, Ltd.**
Kyoto-shi, Kyoto 612-8061 (JP)

(72) Inventors:
- **YOSHIKAWA, Yoshie**
  Kyoto-shi, Kyoto 607-8122 (JP)
- **MUKAI, Hiroyuki**
  Moriyama-shi, Shiga 524-0102 (JP)
- **ASADA, Kiyozo**
  Koga-gun, Shiga 520-3333 (JP)
- **HINO, Fumitsugu**
  Kusatsu-shi, Shiga 525-0028 (JP)
- **KATO, Ikunoshin**
  Uji-shi, Kyoto 611-0028 (JP)

(74) Representative:
VOSSIUS & PARTNER
Postfach 86 07 67
81634 München (DE)

(54) CANCER-ASSOCIATED GENES

(57) A method for detecting cancer cells in an extirpated sample, which comprises monitoring changes in the expression dose of at least one cancer-associated gene selected from among genes whose cDNAs are those hybridizable under stringent conditions with DNAs having the base sequences represented by SEQ ID NOS: 1 to 16 in the Sequence Listing or nucleic acids having the base sequences represented by SEQ ID NOS: 1 to 16 in the Sequence Listing; and a kit, etc., for detecting cancer by the above method.

EP 0 972 830 A1

Description

TECHNICAL FIELD

[0001] The present invention relates to a method for detecting a cancer cell characterized by detecting an expression product of a gene capable of changing an expression level thereof owing to canceration. The present invention relates to a gene capable of changing an expression level thereof and a product of the gene owing to canceration.

BACKGROUND ART

[0002] Cancers constitute the top of the causes for mortality in Japan since 1981, and a gastric cancer occurs especially at the highest frequency. Recently, there has been known that there is a multi-stage carcinogenic mechanism in the course from a normal cell to a cancer cell [Fearon, E.R. et al., *Cell*, **61**, 759 - 767 (1990); Sugimura, T., *Science*, **258**, 603 - 607 (1992)] for which the accumulation of the abnormality in a plurality of genes including DNA repair genes, tumor suppressor genes and oncogenes is essential. Generally, the instability of a gene and the inactivation of a tumor suppressor gene are involved in the development of a cancer, and the activation of an oncogene and/or the overexpression of a growth factor are involved in the advancement and malignancy of a cancer.

[0003] The instability of a gene includes the instability of gene associated with abnormality in a DNA mismatch repair system and the instability at a chromosomal level. An example of the former includes the difference in the chain length of a simple repeated sequence present in a genome between a cancer site and a non-cancer site in the same individual (microsatellite instability) [Thibodeau, S.N. et al., *Science*, **260**, 816 - 819 (1993)], and an example of the latter includes an interchromosomal translocation. The interchromosomal translocation may cause to express a protein which is not found in normal cells, or the interchromosomal translocation may affect an expression level of a protein, even if it is expressed in normal cells. In fact, in human chronic myelocytic leukemia, *bcr* gene is fused with c-*abl* gene by the interchromosomal translocation, and there has been confirmed an expression of a hybrid mRNA transcribed from *bcr-abl* fusion gene, which is absent in normal cells. Further, there has been confirmed that an introduction of *bcr-abl* fusion gene into an animal results in an onset of leukemia [Watson, J.D. et al., *Molecular Biology of Recombinant DNAs*, *2nd Ed.*, Maruzen K. K., 309 (1992)].

[0004] The inactivation of a tumor suppressor gene includes, for example, an inactivation of p53 gene. The inactivation is considered to be caused by a deletion within the gene, or a point mutation occurring in a particular portion of an encoding region [Nigro, J.M. et al., *Nature*, **342**, 705 - 708 (1989); Malkin, D. et al., *Science*, **250**, 1233 - 1238 (1990)]. In addition, since the deletion and the point mutation of the p53 gene are observed in various cancers, and are as frequent as 60% or higher especially in cases of a gastric cancer at an early stage [Yokozaki, H. et al., *Journal of Cancer Research and Clinical Oncology*, **119**, 67 - 70 (1992)], the detection of these mutations is considered to be useful for detecting a cancer at an early stage.

[0005] On the other hand, p16/MTS1 gene has been known to be a gene which is inactivated owing to a homologous deletion, and high-frequency homologous deletions have been observed in cases of a glioma, a pancreatic cancer and a urinary bladder cancer [Cairns, P. et al., *Nature Genetics*, **11**, 210 - 212 (1995)]. p16 Protein regulates a cell cycle, and the abnormality in p16 expression has been suggested to be involved in the canceration of a cell [Okamoto, A. et al., *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 11045 - 11049 (1994)].

[0006] As the causation for the activation of an oncogene, there can be included, for example, a viral insertion mutation in a proximity of an oncogene and an interchromosomal translocation. For example, a viral insertion mutation has been confirmed in lymphoma of a chicken which is caused by an avian leukosis virus (ALV). In this case, it has been found that DNA of an ALV is inserted in the proximity of a gene c-*myc*, and, by potent viral enhancer and promoter, a normal c-*myc* is overexpressed, and a new sequence which is different partially from the normal gene has been expressed. In addition, in a certain kind of human B cell tumor, there has been confirmed that c-*myc*, which is one of oncogenes, is located near a potent transcription signal of immunoglobulin by the interchromosomal translocation, whereby increasing its expression level of the mRNA. In this case, no difference has been found between a protein for c-*myc* in a cancer cell and a protein for c-*myc* expressed in a normal cell, and the canceration is considered to be caused by an increase in the expression level of the c-*myc* mRNA [Watson, J.D. et al., *Molecular Biology of Recombinant DNAs*, *2nd Ed.*, Maruzen K. K., 305 - 308 (1992)].

[0007] An overexpression of a growth factor includes, for example, an overexpression of C-Met which encodes a hepatocyte growth factor receptor. There has been confirmed that the abnormality in expression of the C-Met is observed as an expression of mRNA having the length of 6.0 kb which is not found in a normal mucous membrane at an early stage of gastric cancer [Kuniyasu, H. et al., *International Journal of Cancer*, **55**, 72 - 75 (1993)], or is observed at a high frequency, and that a correlation between the gene amplification and the degree of the cancer malignancy is observed [Kuniyasu, H. et al., *Biochemical and Biophysical Research Communications*, **189**, 227 - 232 (1992)].

[0008] As examples of confirming the correlation between the gene abnormality and the degree of cancer malignancy,

in addition to that of the c-Met mentioned above, there have been confirmed that an amplification and/or an overexpression of an oncogene C-*erbB2* gene is found in mammary cancers, ovarian cancers, gastric cancers and uterine cancers [Wright, C. et al., *Cancer Research*, **49**, 2087 - 2090 (1989); Saffari, B. et al., *Cancer Research*, **55**, 5693 - 5698 (1995)]; and that an amplification and/or an overexpression of an oncogene K-*sam* gene is found in a poorly-differentiated adenocarcinoma which is one tissue type of gastric cancer [Tahara, E. et al., *Gastric Cancer*, Tokyo, Springer-Verlag, Published in 1993, 209 - 217], respectively.

[0009] As described above, the information concerning the gene involved in the development and the advancement of a cancer as well as the abnormality of such genes has been increasing, and the genetic diagnosis of a biopsy material may serve for an early diagnosis and an assessment of the degree of malignancy of a cancer. However, since a carcinogenic mechanism comprises multiple steps and requires an accumulation of a plurality of mutations, a large part of the genes associated with the canceration have still yet been unknown, and further study is necessary. Recently, a gene therapy in which a normal p53 gene is introduced into a cancer cell whereby suppressing the proliferation of the cancer cell is now at a stage of a clinical trial.. Therefore, the solution for a cancer-suppressing gene can shed light not only in the diagnosis but also in the gene therapy.

[0010] Accordingly, a first object of the present invention is to provide a method for detecting cancerated cell and a method for determining a degree of malignancy, on the basis of finding a gene usable as an index for carcinogenesis, particularly a gene capable of changing expression conditions thereof by canceration of a cell, and measuring an expression level of the gene in a resected specimen. A second object of the present invention is to provide a kit used for the above method for detecting a cancer cell and/or a method for determining a degree of malignancy of the cell. A third object of the present invention is to provide a method for controlling proliferation of a cancer cell by using a substance specifically binding to a gene capable of serving as an index for carcinogenesis or an expression product of the gene. Furthermore, a fourth object of the present invention is to provide a novel peptide associated with canceration, and a nucleic acid encoding the peptide.

[0011] To summarize the present invention, a first invention of the present invention is an invention pertaining to a method for detecting a cancer cell in a resected specimen, characterized by determining a change in an expression level of a gene selected from genes of which cDNA corresponds to a DNA comprising a nucleotide sequence as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing, or a DNA capable of hybridizing with a nucleic acid as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing under stringent conditions by, for example, determining a change of an expression level of mRNA or a change of a protein expression level.

[0012] A second invention of the present invention is an invention pertaining to a kit for detecting cancer by the method for detecting of the present invention, characterized in that the kit comprises as an essential constituent any one of primers for amplifying mRNA as an index for a change in an expression level, a probe capable of hybridizing with the above mRNA, or an antibody recognizing a protein as an index for the change in expression level.

[0013] A third invention of the present invention is a method for controlling proliferation of a cancer cell by using a substance specifically binding to the gene or an expression product thereof, characterized in that cDNA of the gene corresponds to a DNA comprising a nucleotide sequence any one of sequences of SEQ ID NOs: 1 to 16 in Sequence Listing, or a DNA capable of hybridizing with DNA as shown in any one of sequences of SEQ ID NOs: 1 to 16 in Sequence Listing, wherein the method gives transcriptional control of the gene and/or functional control of an expression product thereof, and the like.

[0014] A fourth invention of the present invention is an invention pertaining to a peptide usable for detecting cancer and a nucleic acid encoding the peptide, characterized in that the peptide consists of an amino acid sequence comprising an entire portion of an amino acid sequence as shown in SEQ ID NOs: 17 to 19 in Sequence Listing or a partial portion thereof and a nucleic acid encoding the peptide.

[0015] A fifth invention of the present invention pertains to an antibody usable for detecting cancer, the antibody recognizing the above peptide of the fourth invention.

[0016] Incidentally, the term "resected specimen" used in the present specification refers to blood, urine, feces, tissue resected by surgery, and the like. On the other hand, the term "cancer-associated gene" refers to a gene in which the expression conditions thereof change with canceration of a cell.

DISCLOSURE OF INVENTION

[0017] In order to achieve the objects mentioned above, the present inventors have found a cancer-associated gene by comparing the intracellular expression levels of genes between a cancer tissue and a control normal tissue of a cancer patient, and they have found that cancer cells can be detected by comparing the expression level of this gene. In addition, they have found a novel gene in this cancer-associated gene, whereby completing the present invention.

[0018] The terms "cancer tissue" and "control normal tissue" used in the present specification mean a tissue constituting a region of cancerous lesion in a multicellular individual and a tissue constituting a region which is identical spatially to the cancer tissue in the same individual but functions normally.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019]

Figure 1 is an autoradiogram showing electrophoretic patterns of the resulting DNA fragment in a case of detecting cancer-associated genes by DD method.

Figure 2 is an autoradiogram obtained by electrophoresing RNA and then hybridizing a labeled probe with a desired mRNA, in a case of detecting a change in an expression level of mRNA of cancer-associated genes by Northern hybridization method.

Figure 3 is a picture showing electrophoretic patterns of the resulting DNA fragment in a case of detecting a change of expression of a cancer-associated gene by RT-PCR method.

BEST MODE FOR CARRYING OUT THE INVENTION

[0020] The present invention will be explained concretely below.

[0021] The first invention of the present invention provides a method for detecting a cancer cell using an expression level of the cancer-associated gene as an index.

[0022] A gene which can serve as an index for canceration is a gene capable of changing expression conditions thereof by canceration of a cell, namely, a gene of which expression is significantly induced or suppressed. Such a gene can be detected by, for instance, analyzing copy number of the gene on genome or a pattern for translocation in chromosomes, and comparing an expression level of a gene product in a normal cell and a cancerated cell to identify a gene having differences in both cells. The gene product includes, for example, mRNA transcribed from the above gene or a protein which is a translational product. In the detection in the present invention for a cancer-associated gene, it is efficient to use as an index an expression level of mRNA, in which various methods have been developed for its analysis with the progress in gene manipulation technique. Procedures for confirming a change in an expression level of a gene using as an index an expression level of mRNA includes subtractive hybridization method [Zimmermann, C.R. et al., *Cell*, **21**, 709 - 715 (1989)], Representational Difference Analysis (RDA) method [Lisitsyn, N. et al., *Science*, **259**, 946 - 951 (1993)], molecular index method (Japanese Patent Laid-Open No. Hei 8-322598), Differential Display (DD) method [Liang, P. and Pardee, A.B., *Science*, **257**, 967 - 971 (1992)], and the like. Among them, since the procedures of the DD method are simple, the DD method is suitable for screening a gene in the present invention. The method for screening a cancer-associated gene by using the DD method utilized in the present invention will be described in detail below.

[0023] First, mRNA is converted to cDNA by carrying out a reverse transcription reaction with a genome DNA-removed crude RNA sample resulting from treating each RNA individually extracted from a cancer tissue and a control normal tissue to be compared with DNase, together with an oligo(dT) anchor primer and a reverse transcriptase (RTase). Thereafter, the nucleic acid amplification is carried out by polymerase chain reaction (PCR) with the oligo(dT) anchor primer in combination with various random primers.

[0024] Subsequently, a PCR-amplified product obtained separately from the tissues to be compared is subjected to polyacrylamide electrophoresis for each amplified product resulting from a combination of an identical primer pair. The band patterns are compared with each other to find a band exhibiting a difference between the normal cell and the cancer cell. This band is cut out from the gel, and a nucleic acid contained in the band is extracted, whereby a DNA fragment which is considered to be complementary to a partial portion with the mRNA for the cancer-associated gene can be obtained.

[0025] Thereafter, there is studied whether changes in expression levels of mRNA for the cancer-associated gene can be truly confirmed from the DNA fragment obtained in the DD method described above. When the expression level of the mRNA in a normal tissue is confirmed to be higher than that in the cancer tissue, it is determined that the cancer-associated gene is a gene of which expression level is reduced owing to canceration. On the other hand, when the expression level of the mRNA in the cancer tissue is confirmed to be higher than that in the normal tissue, it is determined that the cancer-associated gene is a gene of which expression level is amplified owing to canceration.

[0026] The confirmation on an expression level of mRNA can be made, for example, by labeling the DNA fragment obtained, subjecting a crude RNA sample extracted from either of the cancer tissue or the control normal tissue to Northern hybridization using the above DNA fragment as a detection probe, and confirming the difference in the observed signal intensity with a densitometer. In other words, the stronger the signal intensity, it can be determined that the expression level of the mRNA is high. For example, a signal intensity can be expressed as a value for a volume of a band [IOD (Integrated Optical Density)] obtained from an autoradiogram, or the like. Here, the higher the IOD value, it can be determined that the expression level of the mRNA corresponding to the band is high.

[0027] When the expression level of mRNA is too low so that the change in the expression level of the mRNA cannot be confirmed by means of Northern hybridization analysis, there can be also confirmed with more sensitive RNase pro-

tection assay [Krieg, P.A. and Melton, D.A., *Methods in Enzymology*, **155**, 397 - 415 (1987)] using as a probe RNA prepared from an amplified DNA fragment obtained by the DD method described above, which is derived from mRNA deduced to be expressed from a cancer-associated gene as a template. This method utilizes RNase having substrate specificity wherein it shows cleaving activity on single-stranded RNA, but shows no cleaving activity on double-stranded RNA. Specifically, mRNA to be detected and an excessive amount of a probe are added to a crude RNA sample extracted from a normal tissue and a cancer tissue-derived crude RNA sample, and the mRNA to be detected forms a hybrid with the added probe, whereby acting on an RNase having substrate specificity. The expression level of the mRNA can be confirmed by determining the amount of the double-stranded RNA remaining after the digestion with the RNase mentioned above. In other words, the larger the amount of the remaining double-stranded RNA, it can be determined that the expression level of the mRNA is high.

[0028] The nucleotide sequence of an amplified DNA fragment obtained by the DD method described above, which is derived from mRNA deduced to be expressed from a cancer-associated gene as a template, is sequenced by PCR direct sequencing [Erlich, H.A., *PCR Technology*, Stockton Press, Published in 1989, 45 - 60], or by a combination of TA cloning [Mead, D.A. et al., *Bio/Technology*, **9**, 657 - 663 (1991)] with a usual nucleotide sequencing method to determine the nucleotide sequence, and the amounts of the amplified product as obtained by carrying out RT-PCR with an amplification primer which is designed based on the above nucleotide sequence information are then compared, whereby the mRNA expression level can be confirmed. In other words, the higher the amount of the resulting amplified product, it can be determined that the expression level of the mRNA is high.

[0029] Incidentally, the amplified DNA fragment obtained by the DD method described above, which is derived from mRNA deduced to be expressed from a cancer-associated gene as a template, is not necessarily cDNA complementary to an entire length of mRNA for the cancer-associated gene. In order to obtain cDNA for a cancer-associated gene, for example, a cDNA library derived from a tissue used in screening is prepared; an amplified DNA fragment obtained by the DD method described above, which is derived from mRNA deduced to be expressed from a cancer-associated gene as a template, is labeled; and plaque hybridization is carried out with the labeled cancer-associated gene as a detection probe, whereby cDNA clone for a cancer-associated gene can be isolated.

[0030] The present inventors have succeeded in isolating 14 kinds of DNA fragments comprising a respective nucleotide sequence of a partial portion of cDNA for cancer-associated genes. Genes expressing mRNA which corresponds to cDNA as shown in nucleotide sequences comprising a nucleotide for the DNA fragment thus obtained are named as CA11, CA13, CC24, GG24, AG26, GC31, GC32, GC33, GG33, CC34, GC35, GC36, CA42 and CC62, respectively. Correspondences between SEQ ID NOs in Sequence Listing in which a nucleotide sequence of regions presently determined in each nucleotide sequence of cDNA for 14 kinds of cancer-associated genes and the above name of the gene named by the present inventors are shown in Table 1.

Table 1

SEQ ID NOs in Sequence Listing		Name of Gene
Nucleotide Sequence	Amino Acid Sequence	
1	17	CA11
2	18	CA13
3		CC24
4		GG24
5		AG26
6		GC31
7		GC32
8		GC33
9		GG33
10		CC34
11		GC35
12, 15, 16		GC36
13	19	CA42
14		CC62

[0031] The above cancer-associated genes are roughly classified into a gene in which the expression level is decreased or increased by canceration. The former genes include CA11, AG26, GC35, GC36 and CC62; and the latter genes include CA13, CC24, GG24, GC31, GC32, GC33, GG33, CC34 and CA42.

[0032] By comparing the expression level of each of the genes obtained as above, cancer cells can be detected. In this case, the cancer-associated gene serving as an index may be appropriately selected from the genes listed above, and it may be used as a single kind, or in combination of several kinds of genes. In addition, the cancer-associated gene serving as an index for detection of a cancer cell is not particularly restricted to the 14 kinds of genes listed above, and the cancer-associated gene may be any gene of which cDNA is DNA capable of hybridizing under stringent conditions with the DNA as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing, as long as the expression level of the gene is changed owing to canceration of a cell.

[0033] Conditions capable of hybridizing used in the present specification refer to, for example, those capable of hybridizing by a process comprising incubating DNA immobilized on a nylon membrane with a probe at 65°C for 20 hours in a solution containing 6 × SSC (wherein 1 × SSC is a solution prepared by dissolving sodium chloride 8.76 g and sodium citrate 4.41 g in 1L of water), 1% SDS, 100 μg/ml herring sperm DNA, 0.1% bovine serum albumin, 0.1% polyvinyl pyrrolidone and 0.1% Ficol.

[0034] In fact, there has also been confirmed in the present invention the presence of a gene having the characteristics described above. The nucleotide sequence as shown in SEQ ID NO: 10 in Sequence Listing is present in the nucleotide sequence of cDNA for CC34 gene. DNA as shown in this nucleotide sequence wherein T at base number 935 of the sequence is substituted with A, and 6 bases consisting of the sequence of GTTAAG at a 3'-terminal are deleted has been obtained as a DNA fragment with different amplification levels in the DD method using RNA prepared from a normal tissue and RNA prepared from a cancer tissue. This amplified DNA fragment is capable of hybridizing with DNA as shown in SEQ ID NO: 10 in Sequence Listing. Therefore, a gene expressing mRNA which yields this DNA fragment obtained by the DD method in the present invention is also encompassed in the cancer-associated gene for detecting a cancer cell in the present invention.

[0035] The determination of whether or not a cell is a cancer cell is carried out by firstly using a plurality of normal tissues to confirm a normal level of the expression level of the cancer-associated gene used as an index for canceration by a suitable detection method; subsequently determining an expression level of the cancer-associated gene in a resected specimen; and comparing it with the normal level. Specifically, in a case where the expression of the cancer-associated gene as an index is suppressed by canceration, it is determined to be cancer-positive when the expression of this cancer-associated gene cannot be confirmed or can be confirmed only at a level lower than the normal level in a resected specimen. On the contrary, in a case where the expression of the cancer-associated gene as an index is amplified by canceration, it is determined to be cancer-positive when the expression of this cancer-associated gene is at a level higher than the normal level. In the comparison of the expression level of the cancer-associated gene, there may be employed either the amount of mRNA or the amount of a protein expressed from this gene. Incidentally, the normal level referred in the present specification can be shown by the following equation based on the expression level of the cancer-associated gene in a plurality of normal tissues obtained by an appropriate detection method.

$$[\text{Normal Level Value}] = [\text{Mean Expression Level of Cancer-Associated Gene in Normal Tissue}] \pm 2 \times [\text{Standard Deviation}] \qquad \text{Equation 1}$$

[0036] This normal level value as calculated encompasses 95% of the normal tissues for which the expression level of the cancer-associated gene is determined.

[0037] The detection method utilizing mRNA includes, for example, RT-PCR method, RNase protection assay or Northern hybridization.

[0038] RT-PCR (Reverse transcribed-Polymerase chain reaction) method refers to a method comprising synthesizing cDNA by reverse transcriptional reaction using mRNA as a template, and thereafter performing nucleic acid amplification by PCR [Kawasaki, E.S. et al., *Amplification of RNA. In PCR Protocol, A Guide to Methods And Applications*, Academic Press, Inc., San Diego, 21 - 27 (1991)]. In the present invention, nucleic acid amplification reaction is not particularly limited, and may be Strand Displacement Amplification (SDA) method [Walker, G.T., *Nucleic Acids Res.*, **20**, 1691 - 1696 (1992)], Nucleic Acid Sequence-Based Amplification (NASBA) method [Compton, J., *Nature*, **350**, 91 - 92 (1991)], and the like, in which their reaction conditions are also not particularly limited. In addition, the amplified region of cDNA for the cancer-associated gene is not necessarily an entire length of cDNA, but it may be a partial region of the cDNA, as long as the confirmation of the amplified products is not hindered. It is preferable that a primer pair used in nucleic acid amplification reaction is designed so as to specifically amplify only the cDNA. As long as the confirmation of amplified products for the region is not hindered, it does not matter that cDNA which is not subject to detection may be amplified.

Incidentally, the term "primer" in the present specification refers to an oligonucleotide capable of acting as an initiation

site for DNA synthesis in a case of hybridizing with a template nucleic acid at a suitable temperature under conditions for allowing initiation of synthesis of a primer extension product by DNA polymerase, namely, in the presence of 4 kinds of different nucleotide triphosphates and DNA polymerase in suitable buffer (the buffer being determined by pHs, ionic strength, cofactors, and the like). Typically, the primer comprises 10 to 30 nucleotides. For instance, in a case of CA11 gene in the present specification, there can be exemplified as the former primer a combination of DNAs as shown in SEQ ID NOs: 20 and 21 in Sequence Listing. Hindrance in the confirmation of the amplified products used in the present specification refers, for instance, to a case where the confirmation is carried out by subjecting the amplified DNA fragment to agarose gel electrophoresis, and thereafter staining the gel with ethidium bromide (EtBr), the amount of the amplified DNA fragment present corresponding to mRNA for a cancer-associated gene to be detected cannot be confirmed, since a large number of the amplified DNA fragments having about the same number of bases are produced by nucleic acid amplification reaction, and the separation of each amplified DNA fragment from each other is incomplete.

[0039] Amounts of the amplified DNA level can be confirmed by subjecting the nucleic acid amplification reaction mixture to agarose gel electrophoresis; and confirming from the position and the signal intensity of a band detected with a labeled probe capable of specifically hybridizing with a desired amplified fragment. Therefore, the higher the signal intensity obtained by using a certain amount of a crude RNA sample extracted from a resected specimen, it can be determined that the expression level of a cancer-associated gene to be detected is high. The label on the probe is not particularly limited. For example, there can be used a radioactive substance typically exemplified by ^{32}P, or a fluorescent substance typically exemplified by fluorescein. The signal intensity can, for example, be indicated by IOD of a band on an autoradiogram or a fluorescent image obtained by the method described above.

[0040] On the other hand, when an amplified product can be obtained in a sufficient amount, the amplified product can be confirmed by subjecting it to agarose gel electrophoresis, staining the gel with EtBr, and confirming from the position of the amplified DNA fragment and its fluorescent intensity. Therefore, the higher the fluorescent intensity, it can be determined that the expression level of the cancer-associated gene to be detected is high. It is also possible to determine the expression level of the cancer-associated gene from an IOD of a band on a fluorescent image instead of a fluorescent intensity.

[0041] In order to carry out a more accurate determination, the degree of amplification needs to be expressed numerically. For example, a quantitative PCR method (Japanese Unexamined Patent Publication No. Hei 5-504886) may be applied in the step of nucleic acid amplification reaction, whereby achieving the purpose mentioned above. A typical method includes adding a known amount of a nucleic acid having at its both terminals the primer nucleotide sequences used in amplification of a desired gene and having different internal sequences and sizes as an internal standard and amplifying by PCR reaction; and deducing the desired gene level by comparing the final amplified level of the desired product in the light of the final amplified level of the internal standard. In the present invention, an internal standard is not limited to an externally added standard, and there may also be used cDNA obtained by using as a template mRNA of a gene expressing in a normal tissue and a cancer tissue in the same level. As such cDNA, for example, there can be included cDNA for β-actin gene which is a constituent of a cell backbone.

[0042] For example, in RT-PCR method using a crude RNA sample extracted from gastric cancer tissue cells, when the synthetic oligonucleotides having the nucleotide sequences of SEQ ID NOs: 20 and 21 in Sequence Listing are used as a primer pair for nucleic acid amplification reaction, it is possible to only amplify the nucleotide sequence region as shown in base numbers 46 to 411 in SEQ ID NO: 1 in Sequence Listing of the cDNA nucleotide sequences of a CA11 gene in the present specification as shown in Figure 3(a).

[0043] The expression level of the cancer-associated gene can be determined by RNase protection assay by adding a probe which is RNA in an excess amount capable of specifically hybridizing with mRNA for a cancer-associated gene to be detected or a partial portion thereof to a given amount of a crude RNA sample extracted from a resected specimen, and quantifying the remaining RNA after digestion with the RNase. In other words, the larger the amount of the remaining RNA, it can be determined that the expression level of the cancer-associated gene is high.

[0044] Incidentally, a probe used in this method is not particularly limited, as long as it is RNA capable of hybridizing in hybridization buffer, for example, comprising 80% formamide, 40 mM Pipes (pH 6.4), 400 mM NaCl and 1 mM EDTA at 45°C for 20 hours, and having a nucleotide sequence complementary with a nucleotide sequence specific to mRNA for a cancer-associated gene to be detected. In addition, the label on this probe is not particularly limited, and there may, for example, be used a radioactive substance typically exemplified by ^{32}P, or a fluorescent substance typically exemplified by fluorescein.

[0045] The expression level of the cancer-associated gene can be determined by Northern hybridization by fractionating a given amount of a crude RNA sample extracted from a sample tissue based on the molecular weight; immobilizing on a nylon filter, or the like; bringing mRNA for a cancer-associated gene to be detected into contact with an excess amount of a probe for detecting this gene, and determining the signal intensity obtained from the probe hybridizing with the immobilized RNA. In other words, the higher the signal intensity, it can be determined that the expression level of the cancer-associated gene is high.

[0046] Incidentally, the term "hybridizing" used in the method refers, for example, to those capable of hybridizing by a process comprising incubating at 42°C for 20 hours in hybridization buffer containing 50% formamide, 0.65M NaCl, 0.1M sodium-Pipes, 5 × Denhardt's reagent, 0.1% SDS, 5mM EDTA. The detection probe is preferably a nucleic acid having a nucleotide sequence complementary to a nucleotide sequence which is specific to a cancer associated-gene mRNA to be detected. The nucleic acid is not particularly limited, as long as mRNA to be detected can be particularized by location of the above signals, even if its nucleotide sequence is such that signals can be obtained at several spots in the detection of RNA. Labelling of the above probe is not particularly limited, and there can be used, for example, radioactive substances typically exemplified by ^{32}P, as well as fluorescent substances typically exemplified by fluorescein.

[0047] Figure 2 shows one example of the change in the expression level of mRNA for a cancer-associated gene detected by Northern hybridization method. In this figure, a photograph of an autoradiogram obtained by subjecting each of the RNAs obtained from a cancer tissue and a control normal tissue to electrophoresis individually, and hybridizing with a labeled probe for detecting mRNA for CA11 gene in the present specification.

[0048] In addition, when the change in the expression level of a cancer-associated gene is confirmed using a protein as an index, the confirmation may be made based on the biological activity of the protein, and the detection using an antibody against the protein is preferred for its simplicity in the present invention.

[0049] The antibody in the present invention is an antibody capable of specifically binding to a protein encoded by the cancer-associated gene. Therefore, the larger the amount of the antibody bound to a given amount of a crude protein extracted from a resected specimen, it can be determined that the expression level of the cancer-associated gene is high.

[0050] The protein as an antigen for obtaining the antibody described above can be obtained by purifying from cancer cells expressing the gene, or it can also be obtained by gene engineering technique. For example, a nucleic acid encoding the protein can be obtained by the method described above, in which the DD method is combined with the screening of the cDNA libraries from cells expressing a desired protein. The desired protein can be obtained by incorporating the cDNA obtained into an appropriate expression vector, and expressing it in an appropriate host. Further, this protein may be expressed as a fusion protein. For example, in order to increase the expression level of a desired protein, an appropriate peptide chain is added to the N-terminal or C-terminal derived from other proteins and then allowed to be expressed, and a carrier having an affinity with this peptide chain is used, whereby a desired protein can be purified readily.

[0051] In addition, the antigen for obtaining an antibody may not necessarily be an entire molecule of the protein, and the antigen may be a peptide having an amino acid sequence region which is capable of recognizing the antibody and specific to the protein.

[0052] As the method for obtaining an antibody, the antibody can, for example, be obtained as an anti-serum by immunizing an animal with a peptide together with an adjuvant by a usual method. Alternatively, it can be obtained as a monoclonal antibody according to the method of Galfre, G. et al [Galfre, G. et al., *Nature*, **266**, 550 - 552 (1977)].

[0053] An example of a method for detecting a protein using an antibody includes Western blotting method.

[0054] In this method, the method for detecting with a specific antibody can be carried out by treating cells with a detergent to dissolve intracellular proteins; separating the protein by SDS-polyacrylamide electrophoresis; transferring the resulting protein onto a nitrocellulose membrane, and the like. The antibody bound to a protein can secondarily be detected with, for instance, a ^{25}I-labeled protein A, a peroxidase-linked anti-IgG antibody, and the like.

[0055] The second invention of the present invention provides a kit for detecting a cancer cell. In other words, there can be provided a kit for detecting a cancer cell by utilizing the method for detecting a cancer cell, which is the first invention of the present invention. Concretely, there can be exemplified a kit for detecting the change in the expression level of a cancer-associated gene within the cells using as an index an amount of mRNA or an amount of a protein which is expressed by this gene.

[0056] In the case of a kit for detecting a cancer cell using as an index an expression level of mRNA by using the detection method with the nucleic acid amplification described above in connection with the method for detecting a cancer cell, a primer pair is an essential constituent, where the primer pair has the characteristics described above in connection with the method for detecting a cancer cell wherein the primer pair is capable of detecting mRNA of which cDNA is DNA as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing, or DNA capable of hybridizing under stringent conditions with DNA as shown in a nucleotide sequence comprising the nucleotide sequence as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing. For example, the kit in the present invention utilizing RT-PCR as a detection method may comprise in addition to the primer pair described above reverse transcriptase, dNTPs and a thermostable DNA polymerase. Incidentally, the kinds and the number of the cancer-associated genes to be detected by this kit are not particularly limited. Therefore, the primer pair constituting this kit is not particularly limited, and it may be selected appropriately depending upon the kinds and the number of the cancer-associated genes to be detected.

[0057] One example of the primer pair using as a template cDNA for the cancer-associated gene of the present invention only a part of the region of which is specifically amplified is shown in Table 2. In each primer pair in the table, a symbol of a combination of an alphabet and numerals indicates the name of the primer in the present invention, and a

number within a parenthesis attached to each symbol indicates SEQ ID NO: in Sequence Listing showing the nucleotide sequence of each primer. Incidentally, β-actin shown in Table 2 is a gene selected as an internal standard for the purpose of quantifying mRNA for the cancer-associated gene in a crude RNA sample extracted from a resected specimen.

Table 2

Target Gene	Primer Pair				Size of Amplified DNA Predicted
CA11	F1	(20)	R1	(21)	366 bp
CA13	F2	(22)	R2	(23)	168 bp
CC24	F3	(24)	P3	(25)	259 bp
GG24	F4	(26)	R4	(27)	384 bp
AG26	F5	(28)	R5	(29)	389 bp
GC31	F6	(30)	R6	(31)	213 bp
GC32	F7	(32)	P7	(33)	251 bp
GC33	F8	(34)	R8	(35)	563 bp
GG33	F9	(36)	R9	(37)	218 bp
CC34	F10	(38)	R10	(39)	241 bp
GC35	F11	(40)	R11	(41)	157 bp
GC36	F12	(42)	R12	(43)	95 bp
CA42	F13	(44)	R13	(45)	245 bp
CC62	F14	(46)	R14	(47)	134 bp
β-Actin	F15	(48)	R15	(49)	264 bp

[0058] On the other hand, in the case of a kit for detecting a cancer cell using as an index mRNA by using a detection method employing RNase protection assay or Northern hybridization method, it is an essential requirement for a constituent to have a probe which has the characteristics described above in connection with the method for detecting a cancer and is capable of detecting mRNA of a cancer-associated gene, of which cDNA is DNA comprising the nucleotide sequence as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing, or DNA capable of hybridizing under stringent conditions with DNA as shown in any one of SEQ ID NOs: 1 to 16. For example, in the case of a kit utilizing RNase protection assay, the kit may comprise, in addition to the probe described above, RNase, a concentrated reaction mixture for RNase, and the like. The kinds and the number of the cancer-associated genes to be detected by this kit are not particularly limited. Therefore, a probe constituting this kit is not particularly limited, and it may be selected appropriately depending on the kinds and the number of the cancer-associated genes to be detected.

[0059] On the other hand, in the case of a kit for detecting detection method with an antibody, it is an essential a cancer cell using a protein as an index by using the constituent to have an antibody which has the characteristics described above in connection with the method for detecting a cancer cell and is capable of binding individually and specifically to a peptide encoded by DNA as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing, or DNA as shown in a nucleotide sequence comprising a nucleotide sequence of DNA capable of hybridizing under stringent conditions with DNA as shown in a nucleotide sequence comprising the nucleotide sequence as shown in any one of SEQ ID NOs: 1 to 16. The kinds and the number of the cancer-associated genes to be detected by this kit are not particularly limited. Therefore, the antibody constituting this kit is not particularly limited, and it may be selected appropriately depending upon the kinds and the number of the cancer-associated genes to be detected.

[0060] By using such a kit, a cancer cell can be detected more simply. Therefore, it is possible to diagnose a cancer based on the determined expression level of a cancer-associated gene by using such a kit. In other words, humans whose confirmation of the presence of the cancer cells is made by the method for detecting a cancer cell using this kit can be determined to be cancer-positive.

[0061] The third invention of the present invention is a method for controlling proliferation of a cancer cell using a substance specifically binding to a cancer-associated gene or an expression product thereof. The specific binding substance referred in the present specification can, for example, include nucleic acids, antibodies, cytotoxic T lymphocytes

(CTL), and the like.

[0062] For example, *bcr-abl* chimeric protein detected frequently in chronic myelocytic leukemia has a high tyrosine kinase activity and plays an important role in the onset and the proliferation of the leukemia. An antisense oligonucleotide against a gene encoding this chimeric protein can serve to suppress *in vivo* the proliferation of this gene-expressing tumor (Skorski, T., *Proc. Natl. Acad. Sci. USA* **91**, 4504, 1994). On the other hand, a peptide peculiar to a cancer of a protein expressing specifically in a cancer cell has been conventionally known to be a target of T cell immunoresponse to a cancer cell, and a peptide in a proximal site of the fusion of this fusion protein is immunized, whereby obtaining T cells reactive with this fusion protein (Chen, W., *Proc. Natl. Acad. Sci. USA* **89**, 1468, 1992), which can, for example, be carried out utilizing the techniques described in the following report. Concretely, CD4+T cells which react specifically with a peptide for *ras* in which a 12th amino acid glycine is substituted with another amino acid and which have HLA-DR restrainability are separated in human T cells (Jung, S., *J. Exp. Med.*, **173**, 273, 1991); and from a mouse immunized with a recombinant vaccinia virus capable of producing a protein for *ras* having a mutation in a 61st amino acid a CTL against a peptide consisting of 8 amino acids including such a mutation site can be induced (Skipper, J., *J. Exp. Med.*, **177**, 1493, 1993). Further, in a mouse immunized with a solubilized mutated protein for *ras* prepared by a gene recombination, the proliferation of cancer cells having the same mutation *in vivo* is suppressed (Fenton, R.G., *J. Natl. Cancer Inst.*, **85**, 1294, 1993); and from spleen cells sensitized with a mutated peptide for *ras*, a CTL exhibiting a cytotoxic activity on cancer cells expressing the same mutated ras is obtained (Peace, D.J., *J. Exp. Med.*, **179**, 473, 1994).

[0063] Therefore, as to a gene found to be associated with canceration of cells in the present invention, it is possible to control the cell proliferation by using the same antisense oligonucleotide. In addition, if there can be obtained T cells reactive with a protein encoded by a gene of which expression level is considered to be increased owing to canceration, it is possible to suppress the proliferation of cells in which the protein is expressed at a high level.

[0064] The fourth invention of the present invention provides a novel peptide usable for the detection of cancer, and a nucleic acid encoding the above peptide. In the cancer associated-gene elucidated by the present inventors, genes except for CA11, CA13, GG33, GC35, GC36 and CA42 have been clarified as genes which have already been isolated and identified by homology search with database in which information of nucleotide sequences is recorded. Specifically, CC24 corresponds to cytochrome *c* oxidase subunit I gene [Horai, S. at al., *Proc. Natl. Acad. Sci. USA* **92**, 532 - 536 (1995)]; AG26 corresponds to p190-B gene [Burbelo, P.D. et al., *J. Biol. Chem.* **270**, 30919 - 30926 (1995)]; GC31 corresponds to cytochrome *c* oxidase subunit II gene [Power, M.D. et al., *Nucleic Acids Res.* **17**, 6734 (1989)]; GC32 corresponds to cytochrome *b* gene [Anderson, S. et al., *Nature* **290**, 457 - 465 (1981)]; GC33 corresponds to integrin α 6 subunit gene [Tamura, R.N. et al., *Journal of Cell Biology*, **111**, 1593 - 1604 (1990)]; GG24 corresponds to F1-ATPase β subunit gene [Ohta, S. et al., *The Journal of Biochemistry*, **99**, 135 - 141 (1986)]; and CC62 corresponds to lactoferrin gene [Rey, M.W. et al., *Nucleic Acids Res.* **18**, 5288 (1990)]. On the other hand, CC34 cDNA clone is a clone different from a partial region of the cDNA nucleotide sequence encoding 16SrRNA [Horai, S. et al., *Proc. Natl. Acad. Sci. USA* **92**, 532 - 536 (1995)] by 7 bases. Incidentally, the association with carcinogenesis for these genes has not been known.

[0065] On the other hand, as to each of the genes of CA11, CA13, GG33, GC35, GC36 and CA42, no reports have been yet made with regard to the nucleotide sequence, the sequence identical to the amino acid sequence encoded therein or the sequence having a homology therewith in the region analyzed herein in each of cDNAs for the genes. In other words, in the nucleotide sequence of each of cDNAs for the genes of CA11, CA13, GG33, GC35, GC36 and CA42, a nucleic acid having the nucleotide sequence clarified in the present invention is a novel nucleic acid isolated for the first time by the present inventors.

[0066] As shown in Table 1, a peptide encoded by a novel nucleic acid in the present invention comprising the nucleotide sequence as shown in each of SEQ ID NOs: 1, 2 and 13 in Sequence Listing is deduced based on this nucleotide sequence that the peptide comprises the amino acid sequence as shown in each of SEQ ID NOs: 17, 18 and 19 in Sequence Listing, without being limited thereto. Specifically, there also are encompassed [1] a peptide comprising an entire portion of the amino acid sequence as shown in any one of SEQ ID NOs: 17 to 19 in Sequence Listing, or a partial portion thereof; and [2] a peptide resulting from addition, deletion or substitution of one or more amino acids in the amino acid sequence as shown in any one of SEQ ID NOs: 17 to 19 in Sequence Listing, and having a change in the expression level owing to canceration of cells, because of the reasons described below.

[0067] In a naturally-occurring protein, mutations such as deletion, insertion, addition and substitution of amino acids can take place in its amino acid sequence in addition to a polymorphism or a mutation in a gene encoding it as well as a modification *in vivo* or in purification step after its production. Nevertheless, when such a mutation is present in a region in which it is insignificant to preserve the activities and the structure of the protein, there have been known to exhibit physiological and biological activities substantially of the same level as those of the proteins without mutations.

[0068] In addition, the same can be said for the case where the mutations described above are artificially introduced into an amino acid sequence of the protein, in which case diversified, various kinds of mutants can be further prepared. For instance, it has been also known that a polypeptide resulting from substitution of a particular cysteine residue with serine in the amino acid sequence of human interleukin 2 (IL-2) retains IL-2 activity [Wang, A. et. al., *Science*, **224**,

1431 - 1433 (1984)]. Therefore, proteins are encompassed within the scope of the present invention, as long as no difference in the change in an expression level owing to canceration is found, even if the protein has an amino acid sequence which results from deletion, insertion, addition or substitution of one or several amino acid residues in an amino acid sequence disclosed by the present invention.

[0069] Further, certain kinds of proteins have been known to have a peptide region which is unessential for its activity. Examples are signal peptide present in a protein secreted extracellularly, and a pro-sequence found in a precursor of a protease, or the like, and almost all of these regions are removed after translation or when converted into an active protein. Such proteins are present in the form of different primary structures, but the proteins exhibit equivalent functions eventually.

[0070] When a protein is produced by a gene engineering technique, a peptide chain irrelevant to the activity of a desired protein may be added to an amino terminal or carboxyl terminal of the protein. For example, in order to increase the expression level of a desired protein, a fusion protein resulting from adding a part of an amino terminal region of a protein highly expressed in a host used to an amino terminal of a desired protein may be prepared. Alternatively, in order to facilitate the purification of the protein expressed, a peptide having an affinity with a particular substance may be added to an amino terminal or carboxyl terminal of a desired protein. These added peptides may remain in an added state when there is no adverse effect on the activity of a desired protein, or the added peptides may be removed from a desired protein, if necessary, by means of an appropriate treatment such as a limited degradation with a protease.

[0071] Even a protein having or adding a peptide unessential for its function is also encompassed within the scope of the protein of the present invention, as long as it can exhibit an equivalent function. The term "peptide" in the present specification refers to two or more amino acids bound to each other via peptide bonds, and is intended to encompass those referred to as "protein."

[0072] A partial portion of the novel nucleic acid in the present invention consists of a nucleic acid encoding a peptide having the amino acid sequence as shown in any one of SEQ ID NOs: 17 to 19 in Sequence Listing, wherein its nucleotide sequence include those as shown in Table 1, for instance, the nucleotide sequence as shown in any one of SEQ ID NOs: 1, 2 and 13 in Sequence Listing. In other words, the peptide having the amino acid sequence as shown in SEQ ID NO: 17 in Sequence Listing is encoded by the base numbers 3 to 584 of the nucleotide sequence as shown in SEQ ID NO: 1 in Sequence Listing; the peptide having the amino acid sequence as shown in SEQ ID NO: 18 in Sequence Listing is encoded by the base numbers 1698 to 1850 of the nucleotide sequence as shown in SEQ ID NO: 2 in Sequence Listing; the peptide having the amino acid sequence as shown in SEQ ID NO: 19 in Sequence Listing is encoded by the base numbers 8 to 196 of the nucleotide sequence as shown in SEQ ID NO: 13 in Sequence Listing, respectively, but the nucleic acids encoding the novel peptide in the present invention are not limited thereto. Specifically, there are also encompassed within the present invention 1) a nucleic acid encoding a peptide usable for detection of a cancer cell, wherein the peptide comprises an entire sequence of the amino acid sequence as shown in any one of SEQ ID NOs: 17 to 19 in Sequence Listing, or a partial sequence thereof; 2) a nucleic acid encoding a peptide capable of changing its expression level owing to canceration of a cell, wherein the nucleic acid is capable of hybridizing with the novel nucleic acid of the present invention under stringent conditions; 3) a nucleic acid encoding a peptide usable for detection of a cancer cell by the change in its expression level, wherein one or more amino acids are added, deleted or substituted in the amino acid sequence as shown in any one of SEQ ID NOs: 17 to 19 in Sequence Listing, and the like.

[0073] The term "nucleic acid encoding an amino acid sequence" described in the present specification will be described. There has been known that as the codon (triplet base combination) designating a particular amino acid on a gene, 1 to 6 kinds each exist for every amino acid. Therefore, there can be a large number of nucleic acids each encoding an amino acid sequence, depending on its amino acid sequence. In nature, the gene does not exist in a stable form, and it is not rare that the mutation of its nucleotide sequence takes place. The mutation on the gene may not affect the amino acid sequence to be encoded (so-called "silent mutation"), in which case it can be said that different nucleic acids encoding the same amino acid sequence have been produced. There cannot, therefore, be denied the possibility that even when the nucleic acid encoding a particular amino acid sequence is isolated, a variety of nucleic acids encoding the same amino acid sequence are produced with generation passage of the organism containing them. Moreover, it is not difficult to artificially produce a variety of the nucleic acids encoding the same amino acid sequence by means of various genetic engineering techniques. For example, when a codon used on a natural nucleic acid encoding the desired protein is low in usage in the host in the production of a protein by genetic engineering, the expression level of the protein is sometimes insufficient. In such a case, high expression of the desired protein is achieved by artificially converting the codon into another one of commonly used in the host without changing the amino acid sequence encoded (for example, Japanese Examined Patent Publication No. Hei 7-102146). It is of course possible to artificially produce a variety of nucleic acids encoding a particular amino acid sequence, and the nucleic acids can be also produced in nature. Therefore, the present invention includes a nucleic acid, as long as the nucleic acid encodes an amino acid sequence disclosed in the present specification, even if it is not a nucleic acid having same nucleotide sequence disclosed in the present specification.

[0074] In fact, in the present invention, nucleic acids of which nucleotide sequences are slightly different but the amino acid sequence encoded is identical is obtained. Although R at base number 1784 is A, and K at base number 1985 is T in the nucleotide sequence as shown in SEQ ID NO: 2 in Sequence Listing of which the nucleotide sequence is contained in a nucleotide sequence for cDNA of CA13 gene, there is obtained cDNA in which R at base number 1784 is G, and K at base number 1985 is T; and a nucleic acid in which R at base number 1784 is A, and K at base number 1985 is G in the nucleotide sequence as shown in SEQ ID NO: 2 in Sequence Listing. However, the differences of the nucleotide sequence at these two sites do not affect the amino acid sequence encoded in base numbers 1698 to 1850 in the nucleotide sequence as shown in SEQ ID NO: 2 in Sequence Listing, and each peptide encoded by the above three kinds of nucleic acids has the amino acid sequence as shown in SEQ ID NO: 18 in Sequence Listing.

[0075] Further, cDNAs for GG33, GC35 as well as GC36 genes in the novel genes of the present invention respectively have the nucleotide sequence as shown in SEQ ID NOs: 9, 11, as well as 12, 15 and 16 in Sequence Listing.

[0076] Moreover, the novel nucleic acids of the present invention include a nucleic acid capable of hybridizing with the nucleic acid having the nucleotide sequences as shown in any one of SEQ ID NOs: 9, 11 as well as 12, 15 and 16 in Sequence Listing under stringent conditions, wherein the nucleic acid is complementary to a nucleotide sequence for mRNA capable of changing an expression level by canceration. In fact, the nucleic acid having the above properties is obtained in the present invention. For instance, there are obtained the above nucleic acid of which nucleotide sequence is slightly different but an encoded amino acid sequence is identical; or of which nucleotide sequence has the nucleotide sequence resulting from the substitution of T at base number 935 with A and the deletion of 6 bases consisting of sequence GTTAAG at 3' terminal in the nucleotide sequence as shown in SEQ ID NO: 10 in Sequence Listing, wherein cDNA of CC34 gene, which is not a novel nucleic acid in the present invention, comprises the nucleotide sequence as shown in SEQ ID NO: 10 in Sequence Listing.

[0077] In addition, the fifth invention of the present invention provides an antibody against the peptide encoded by the novel nucleic acid in the present invention. The above antibody can be utilized for detection of the cancer cell described above.

[0078] The present invention will be described more concretely hereinbelow by means of the working examples, without intending to restrict the scope of the present invention thereto.

Example 1 Analysis of Cancer-Associated Gene

1) Confirmation of mRNA Which Can Serve As Index for Detecting Cancer

[0079] There was confirmed whether or not mRNA of which expression level was changed by canceration was present by DD method comprising comparing the expression of mRNA of a cancerated lesion tissue with that of a control normal tissue of a stomach as detailed below.

[0080] First, from each of a cancer tissue and a control normal tissue of a stomach excised from a patient with an advanced, poorly-differentiated adenocarcinoma, RNA was extracted with TRIzol™ reagent (manufactured by Gibco BRL) to obtain a crude RNA sample. A 50 μg portion of the crude RNA sample thus obtained was reacted with 10 units of DNaseI (manufactured by Takara Shuzo Co., Ltd.) at 37°C for 30 minutes in the presence of 5 mM $MgCl_2$ as a final concentration and 20 units of RNase inhibitor (manufactured by Takara Shuzo Co., Ltd.) to remove genomic DNA. Using this RNA, RT-PCR was carried out with Differential Display™ Kit (manufactured by Display Systems) and Enzyme Set-DD (manufactured by Takara Shuzo Co., Ltd.) in accordance with the procedures described in the instruction attached to the kit.

[0081] Specifically, reverse transcription reaction was carried out per one reaction by mixing 200 ng of the crude RNA sample treated with the above DNase, and any one kind of the oligonucleotides having the nucleotide sequences as shown in SEQ ID NOs: 56 to 64 in Sequence Listing as a primer, thereafter heat-treating at 70°C for 10 minutes, subjecting to rapid cooling, and subsequently reacting with AMV reverse transcriptase at 55°C for 30 minutes. Other downstream primers were individually reacted in the same manner to prepare 9 kinds of single-stranded cDNA samples in total.

[0082] In the subsequent nucleic acid amplification reaction by PCR, a nucleic acid amplification was carried out by PCR using each of the 9 kinds of single-stranded cDNAs described above as a template, an oligo(dT) primer identical to that used in the reverse transcription as a downstream primer, and any one kind of the 10mer-oligonucleotides in the kit which had the nucleotide sequences as shown in SEQ ID NOs: 50 to 55 in Sequence Listing as an upstream primer, to prepare 54 kinds of amplified DNA samples in total.

[0083] The PCR was carried out by adding 3 mM $MgCl_2$, 15 μM each of dATP, dGTP, dCTP and dTTP as substrates, and 1.85 kBq/ml [α-^{33}P]-dATP (manufactured by Amersham) as a labelling compound, and reacting for 40 cycles, wherein one cycle consists of at 94°C for 30 seconds, at 40°C for 60 seconds and at 72°C for 60 seconds. After termination of the reaction, an equivolume of 95% formamide was added, and the mixture was subjected to thermal denaturation at 90°C for 2 minutes to obtain a sample for electrophoresis. The electrophoresis was carried out on a 7 M urea-

denatured 5% polyacrylamide gel, and autoradiography yielded a fingerprint comprising a large number of bands, wherein there were found to be bands having different signal intensities between the autoradiogram of the cancer tissue and that of the control normal tissue.

[0084] As one example, the results where D4 having the nucleotide sequence as shown in SEQ ID NO: 59 in Sequence Listing was used as a downstream primer, and U1 having the nucleotide sequence as shown in SEQ ID NO: 50 was used as an upstream primer are shown in Figure 1. Specifically, Figure 1 is a reproduced photograph of an autoradiogram showing electrophoretic patterns of the DNA fragment obtained when a cancer-associated gene was detected by the DD method. Here, in Figure 1, 1N is a lane wherein on an acrylamide gel was electrophoresed an amplified DNA fragment obtained by using as a template a crude RNA sample obtained from a normal tissue of a patient with a poorly-differentiated adenocarcinoma-type gastric cancer; and 1T is a lane wherein on an acrylamide gel was electrophoresed an amplified DNA fragment obtained by using as a template a crude RNA sample obtained from a cancer tissue of the same patient with the poorly-differentiated adenocarcinoma-type gastric cancer, respectively. A band having a stronger signal intensity in the autoradiogram obtained from the control normal tissue than in the autoradiogram of the cancer tissue sample was found at the position corresponding to about 750 bp as indicated with "→" in Figure 1. The present inventors named the gene expressing the mRNA which causes the band to show this difference in the intensity as CA11.

[0085] Table 3 showed the combination of the upstream and downstream primers for detecting the difference in the expression level of each mRNAs by means of the DD method, an the approximate size of an amplified DNA fragment, and the difference in the level of the amplified DNA obtained by RT-PCR from the cancer tissue and the control normal tissue for each of genes which was detected by the present inventors with the DD method and named. In the column of the primers in Table 3, a symbol of a combination of an alphabet and numerals indicates the name of a primer, and a number within a parenthesis attached to each symbol indicates SEQ ID NO: showing the nucleotide sequence of the primer in Sequence Listing.

Table 3

Name of Gene	Primer Pair		Approximate Size of Amplified DNA fragment	Difference in Amount of DNA fragment
	Upstream	Downstream		
CA11	U1 (50)	D4 (59)	750 bp	Cancer Tissue < Normal Tissue
CA13	U1 (50)	D4 (59)	620 bp	Cancer Tissue > Normal Tissue
CC24	U2 (51)	D5 (60)	800 bp	Cancer Tissue > Normal Tissue
GG24	U2 (51)	D9 (64)	480 bp	Cancer Tissue > Normal Tissue
AG26	U2 (51)	D3 (58)	550 bp	Cancer Tissue < Normal Tissue
GC31	U3 (52)	D8 (63)	440 bp	Cancer Tissue > Normal Tissue
GC32	U3 (52)	D8 (63)	310 bp	Cancer Tissue > Normal Tissue
GC33	U3 (52)	D8 (63)	300 bp	Cancer Tissue > Normal Tissue
GG33	U3 (52)	D9 (64)	410 bp	Cancer Tissue > Normal Tissue
CC34	U3 (52)	D5 (60)	290 bp	Cancer Tissue > Normal Tissue
GC35	U3 (52)	D8 (63)	210 bp	Cancer Tissue < Normal Tissue
GC36	U3 (52)	D8 (63)	190 bp	Cancer Tissue < Normal Tissue
CA42	U4 (53)	D4 (59)	660 bp	Cancer Tissue > Normal Tissue
CC62	U6 (55)	D5 (60)	380 bp	Cancer Tissue < Normal Tissue

2) Identification of mRNA Serving as Index for Detecting Cancer

[0086] There was investigated whether a change in an expression level of the mRNA used as a template for an amplified DNA fragment derived from each of the genes shown in Table 3 as confirmed by the DD method in Section 1) described above was truly associated with canceration.

[0087] First, the studies were made by means of Northern hybridization. Specifically, there was studied whether the

difference in the expression levels of the mRNA of a cancer-associated gene expressed in a cancer tissue and that in a control normal tissue could be detected by using each amplified DNA fragment obtained by the method in Section 1) described above as a probe.

[0088] The probe for the detection was prepared as follows. Specifically, from-the acrylamide gel on which the amplified DNA fragment obtained by the DD method in Section 1) described above was electrophoresed, the region containing each amplified DNA fragment shown in Table 3 was cut out, and thereto was added 100 μl of water and subjected to a heat extraction to collect individually each DNA fragment contained. Re-amplification by PCR was carried out by using each DNA fragment individually as a template, with a combination of the upstream and downstream primers used to obtain each DNA fragment shown in Table 3. Further, about 100 ng of each amplified DNA fragment was labeled with ^{32}P using Random Primer DNA Labeling Kit (manufactured by Takara Shuzo Co., Ltd.) to prepare 14 kinds of probes for detection. Separately from above, mRNA for β-actin gene was selected as a positive control of a crude RNA extracted from each tissue, and the synthetic oligonucleotide having the nucleotide sequence as shown in SEQ ID NO: 65 in Sequence Listing was labeled in the same manner with ^{32}P to obtain a probe for detecting mRNA for β-actin gene. Thereafter, the probe for detection described above was mixed together with herring sperm DNA so as to have a concentration of 100 μg/ml, and then heat-denatured. To the resulting reaction mixture was added hybridization buffer (50% formamide, 0.65 M NaCl, 0.1M Na-Pipes, 5 × Denhardt's reagent, 0.1% SDS, 5 mM EDTA) to prepare 15 kinds of probe solutions for detection in Northern hybridization.

[0089] Northern hybridization was carried out as follows. First, 20 μg per well of a crude RNA sample extracted from each of a cancer tissue and a control normal tissue from the patient with a gastric cancer prepared as described above was subjected to electrophoresis on a formalin-denatured 1% agarose gel and blotted on a Hybond N^+ membrane (manufactured by Amersham). Subsequently, a blotted membrane and hybridization buffer added with heat-denatured herring sperm DNA so as to have final concentration of 100 μg/ml were added to a Hybri Bag (manufactured by COSMO BIO). The resulting composition was allowed to stand at 42°C for 2 hours, and then the buffer was discarded to prepare a membrane with pre-hybridization treatment. After preparing 15 such membranes as above, to each membrane was added each of the 15 kinds of detection probe solutions for Northern hybridization described above, and hybridization was carried out at 42°C for 16 hours. Thereafter, each blotted membrane was taken from the Hybri Bag, and rinsed with washing solution I (2 × SSC, 0.2% sodium pyrophosphate, 0.1% SDS) at 42°C for 20 minutes, and then with washing solution II (0.5 × SSC, 0.2% sodium pyrophosphate, 0.1% SDS) at 42°C for 20 minutes. Incidentally, rinsing with washing solution II was repeated twice with replacing the washing solution. The membrane after rinsing was wrapped with a plastic film and exposed for one day and night to a high-sensitivity X-ray film (manufactured by Kodak). From the signal intensity in the resultant autoradiogram, the expression level in the cancer tissue was compared with that of the control normal tissue.

[0090] As one example, the results of the detection of mRNA for CA11 gene are shown in Figure 2. In Figure 2, 1N is a lane wherein on an agarose gel was electrophoresed a crude RNA sample obtained from a normal tissue of a patient with a poorly-differentiated adenocarcinoma-type gastric cancer; and 1T is a lane wherein on an agarose gel was electrophoresed a crude RNA sample obtained from a cancer tissue of the same patient with the poorly-differentiated adenocarcinoma-type gastric cancer. (a) shows results obtained with a probe for detecting CA11, and (b) shows results obtained with a probe for detecting β-actin. Since both of the 1N and the 1T exhibited the signals obtained with the probes for detecting β-actin as shown in (b), it is clear that in the both samples the RNA is extracted without undergoing degradation excessively. On the other hand, a clear signal as indicated by "→" at a position near 1.1 kb was present only in lane 1N but no signals were present in lane 1T as shown in (a). Therefore, it was found that the CA11 was a gene of which expression level was reduced owing to canceration. Similarly, CC62 exhibited a band at about 2.6 kb only on the autoradiogram derived from the control normal stomach tissue. GC31, GC32 and CC34 showed the bands at about 1.0 kb, about 1.6 kb and about 1.7 kb, respectively, and in any of these genes more intensive signal was obtained for the crude RNA samples prepared from the gastric cancer tissues as compared to that of the crude RNA samples prepared from the control normal stomach tissues. Incidentally, the signal intensity was determined by measuring each band of an autoradiogram by a densitometer. Subsequently, IOD of each band obtained on the autoradiogram was calculated with FMBIO-100 (manufactured by Hitachi Soft Engineering), and an index was calculated by the equation as shown below to determine whether or not a gene was a cancer-associated gene.

$$[\text{Index Value}] = (X \times \beta Y)/(Y \times \beta X) \qquad \text{Equation 2:}$$

[0091] In the above equation, each symbol expresses the following value:

X: IOD of a band derived from mRNA for the gene shown in Table 3 obtained from a gastric cancer tissue;

Y: IOD of a band derived from mRNA for the gene shown in Table 3 obtained from a control normal stomach tissue;

βX: IOD of a band derived from mRNA for β-actin gene obtained from a gastric cancer tissue; and

βY: IOD of a band derived from mRNA for β-actin gene obtained from a control normal stomach tissue.

[0092] The comparison on the expression level was made by carrying out RT-PCR with respect to each of the genes CA13, CC24, GG24, AG26, GC33, GG33, GC35, GC36 and CA42 in which no signals were obtained by Northern hybridization. In order to design a primer for the nucleic acid amplification reaction in the RT-PCR, each DNA fragment used as a probe in Northern hybridization was subjected to a direct sequencing by PCR, or was cloned by a TA cloning procedure and then sequenced by a dideoxy method, whereby determining its nucleotide sequence. The nucleotide sequences of primers designed based on the resulting nucleotide sequence information and used in the RT-PCR with mRNA derived from each of the genes as a template are as shown in any of SEQ ID NOs: 22 to 29, 34 to 37 and 38 to 43 in Sequence Listing. Table 2 shows the genes together with the corresponding primers used to confirm their expression.

[0093] A change in an expression level of mRNA by RT-PCR was confirmed by a DNaseI treatment of a crude RNA sample obtained from each of the cancer tissue and the control normal tissue of a patient with a gastric cancer prepared by the method in Section 1) described above. Thereafter, RT-PCR was carried out in a 100 μl reaction system of 40 ng of each treated sample with TaKaRa RNA PCR Kit Ver. 2.1 according to the procedures described in the instruction attached to the kit. Specifically, 40 ng of a crude RNA sample as a template and an oligo(dT) primer (final concentration: 2.5 μM) as a downstream primer were used to prepare a reverse transcription reaction mixture (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 5 mM $MgCl_2$, 1 mM each of dNTPs, 100 units of RNase inhibitor, 25 units of AMV reverse transcriptase), and the reverse transcription reaction was carried out at 30°C for 10 minutes, and at 55°C for 20 minutes and then at 95°C for 5 minutes. Each 10 μl of the reverse transcription reaction mixture was added to each 40 μl of 10 kinds of PCR reaction mixtures (final concentration: 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 1.25 units of TaKaRa Taq DNA polymerase) individually containing the primer pairs (0.2 μM) for detecting each of the mRNAs for the genes of CA13, CC24, GG24, AG26, GC33, GG33, GC35, GC36, CA42 and β-actin to make up a volume of 50 μl. One cycle after the pre-incubation at 94°C for 2 minutes in PCR consisted of the step of incubation, at 94°C for 30 seconds, at 55°C for 60 seconds, and then at 72°C for 60 seconds. The amount of an amplified DNA product was quantified by subjecting the amplified DNA product to agarose gel electrophoresis, staining the gel with ethidium bromide, calculating the IOD of each band on the fluorescent image with FMBIO-100 to obtain an index for determining whether or not a gene is a cancer-associated gene from Equation 2 shown above.

[0094] The results of Northern hybridization method and RT-PCR described above, and the patterns of the changes in the expression owing to the canceration of each of the genes evident from these results were shown in Table 4. In the column of the patterns of the changes in the expression, a gene of which expression was amplified owing to canceration was indicated with "↑" and a gene of which expression was suppressed owing to canceration was indicated with "↓". Specifically, it was determined in Table 4 that a gene having an index value greater than 1 is a gene of which expression level was increased owing to canceration, and a gene having an index value less than 1 is a gene of which expression level was reduced owing to canceration. As a result, there were clarified that the genes CA13, CC24, GG24, GC31, GC32, GC33, GG33, CC34 and CA42 were those of which expression levels were increased owing to canceration, and the genes CA11, AG26, GC35, GC36 and CC62 were those of which expression levels were reduced owing to canceration.

Table 4

Name of Gene	Index Value	Method for Determining Index Value	Patterns of Changes in Expression
CA11	0.036	A	↓
CA13	6.3	B	↑
CC24	2.0	B	↑
GG24	2.8	B	↑
AG26	0.52	B	↓
GC31	3.1	A	↑
GC32	3.6	A	↑
GC33	2.3	B	↑
GG33	2.2	B	↑
CC34	15	A	↑

Table 4 (continued)

Name of Gene	Index Value	Method for Determining Index Value	Patterns of Changes in Expression
GC35	0.0046	B	↓
GC36	0.048	B	↓
CA42	1.9	B	↑
CC62	0.56	A	↓

(note) In the table, "A" represents a determination from the autoradiogram in Northern hybridization, and "B" represents a determination based on the electrophoretic photograph of the amplified product by RT-PCR.

3) Acquisition of Cancer-Associated Gene cDNA

[0095] A cDNA fragment of each of these cancer-associated genes was then cloned. First, a cDNA library was prepared by fractionating mRNA from a crude RNA sample derived from a cancer tissue or a normal tissue, which was prepared by the method described in Section 1) with mRNA Purification Kit (manufactured by Pharmacia) on an oligo(dT) column, and plating a phage and a host cell XLI-Blue MRF' at a cell density of about 40,000 plaques per rectangular plate in a 10 cm × 14 cm plate using a ZAP-cDNA synthesis kit (manufactured by Stratagene) according to the protocols attached to the kit. Thereafter, phage particles were transferred onto a Hybond N^+ membrane, and screening was carried out by means of plaque hybridization using a probe identical to that used in Northern hybridization described in Section 2), whereby finding a Uni-ZAP XR clone containing a desired cDNA gene. This recombinant Uni-ZAP XR clone was converted into a pBluescript phagemide by means of an *in vitro* excision method. The nucleotide sequence of a DNA fragment incorporated into this recombinant phagemide was determined with a fluorescent DNA sequencer (manufactured by ABI). The nucleotide sequences obtained from connection of the nucleotide sequences of the cDNA fragments contained in the cDNA library by means of walking based on the nucleotide sequence of the DNA fragment incorporated into the phagemide are shown in SEQ ID NOs: 1 to 14 in Sequence Listing. Incidentally, a cDNA clone of about 2.6 kbp was obtained regarding the gene GC36. As a result of analyzing the nucleotide sequences from its 5'-terminal and 3'-terminal of the cDNA clone, the nucleotide sequence information as shown in SEQ ID NO: 15 and SEQ ID NO: 16 in Sequence Listing was obtained.

[0096] Each of the nucleotide sequences thus obtained was subjected to a homology search with known gene cDNA nucleotide sequences recorded in Genebank by using BLAST program [Altschul, S.F., *Journal of Molecular Biology*, **215**, 403 - 410, (1990)]. As a result, there have not been reported any sequences corresponding to the cDNA of each of CA11, CA13, GG33, GC35, GC36 and CA42, so that these genes were determined to be novel genes. Further, as a result of searching an open reading frame for a gene product based on the nucleotide sequence contained in each of the gene cDNAs of CA11, CA13 and CA42, it was deduced that CA11cDNA encodes the amino acid sequence as shown in SEQ ID NO: 17 in Sequence Listing, CA13cDNA encodes the amino acid sequence as shown in SEQ ID NO: 18 in Sequence Listing, and CA42cDNA encodes the amino acid sequence as shown in SEQ ID NO: 19 in Sequence Listing, respectively. On the other hand, CC24 corresponded to cytochrome *c* oxidase subunit I gene, AG26 to p190-B gene, GC31 to cytochrome *c* oxidase subunit II gene, GC32 to cytochrome *b* gene, GC33 to integrin α 6 subunit gene, GG24 to F1-ATPase β subunit gene, and CC62 to lactoferrin gene. Moreover, the nucleotide sequence region as shown in SEQ ID NO: 10 in Sequence Listing for the CC34 cDNA was found to be different from a partial region of the cDNA encoding a mitochondrial 16SrRNA by 7 bases.

[0097] Incidentally, in the screening of the cDNA library using as a probe an amplified DNA fragment derived from CC34, in addition to the cDNA clone having the nucleotide sequence as shown in SEQ ID NO: 10 in Sequence Listing, an additional, different kind of positive cDNA clone was obtained. There was clarified that the nucleotide sequence of this cDNA had a nucleotide sequence in which T at base number 935 in the nucleotide as shown in SEQ ID NO: 10 in Sequence Listing was substituted with A, and 6 bases consisting of GTTAAG at the 3'-terminal were deleted, of which 1540 bases out of the entire 1546 bases of the entire nucleotide sequence had an identical sequence to a partial region of the cDNA encoding a mitochondrial 16SrRNA.

Example 2 Confirmation of Change in Gene Expression in Cancer Tissue

[0098] With respect to each cancer-associated gene confirmed in Example 1, the association of the expression of this gene with the canceration of cells was evaluated by using a cancer tissue different from that used in Example 1.

1) Confirmation of Change in Gene Expression in Cancer Tissue of Patient with Signet Ring Cell Gastric Cancer

[0099] Using a crude RNA sample prepared in the same manner as in Section 1) of Example 1 from each of a cancer tissue and a control normal tissue excised from a patient with a signet ring cell gastric cancer who was different from the one provided the tissues used in Sections 1) and 2) of Example 1, the expression levels in the cancer tissue and the normal tissue were compared with respect to each of the 14 kinds of cancer-associated genes clarified in Section 3) of Example 1 by using the expression level of the mRNA as an index by means of carrying out Northern hybridization or RT-PCR described in Section 2) of Example 1. As one example, the results of the detection of mRNA for CA11 gene by RT-PCR method are shown in Figure 3. Specifically, Figure 3 is a photograph of a fluorescent image of the electrophoresis of a DNA fragment obtained when a change in an expression level of a cancer-associated gene is detected by RT-PCR method. The reaction conditions of the RT-PCR were according to the method described in Section 2) of Example 1, with setting two patterns in the number of the cycles of the PCR, i.e., 25 and 30. In Figure 3, (a) shows the results of the detection of the expression of a cancer-associated gene CA11, and (b) shows the results of the confirmation of the expression of β-actin as a positive control. In Figure 3, 2T is an amplified DNA fragment obtained by using as a template a crude RNA sample extracted from a gastric cancer tissue of the patient with a signet ring cell gastric cancer, and 2N is an amplified DNA fragment obtained by using as a template a crude RNA sample extracted from a normal gastric tissue of the patient with the signet ring cell gastric cancer. Also, the numerals "25" and "30" in Figure 3 are the numbers of the cycles of the nucleic acid amplification in the RT-PCR method. Table 5 shows the results of calculated IODs of the bands on the fluorescent image shown in Figure 3. Incidentally, each index shown in Table 5 was calculated from Equation 2 described in Section 2) of Example 1.

Table 5

Number of Cycles	25		30	
Sample Name	2T	2N	2T	2N
CA11	365	31118	6345	61742
β-Actin	710	562	25115	20425
Index Value	0.0093		0.083	

[0100] In Table 5, since the IOD values of the band derived from β-actin obtained on the fluorescent image of 2T and 2N were of the similar level in the PCR cycles of 25 and 30, there was clarified that RNAs could be extracted from all samples. However, since the index was less than 1 for both the 25 and 30 cycles of the PCR, there was clarified that CA11 was a gene of which expression level was reduced owing to canceration even also with patients with a signet ring cell gastric cancer. With respect to 13 kinds of cancer-associated genes other than CA11, there was found to be a change in the expression level in the same manner as in Section 2) of Example 1, so that there was clarified that the change in the expression level of each of the 14 kinds of genes as clarified in Section 3) of Example 1 was not a change peculiar to the tissue of the patient tested in Section 1) of Example 1.

Example 3

Construction of Kit for Detecting Cancer

[0101] A kit for detecting a cancer utilizing RT-PCR method comprising the following components was constructed.

[0102] Specifically, a kit comprises DNaseI, AMV reverse transcriptase, RNase inhibitor, 10 × RT-PCR buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl), 25 mM $MgCl_2$, and a mixture of 2.5 mM each of dATP, dGTP, dCTP and dTTP, an oligo(dT) primer, Taq DNA polymerase, a primer pair specific to each of the genes and a primer pair for amplifying β-actin gene as a positive control shown in Table 2. In the column of the primer pair in Table 2, a symbol of a combination of an alphabet and a numeral indicates the name of a primer, and a number within a parenthesis following each symbol indicates SEQ ID NO: showing the nucleotide sequence of the primer in Sequence Listing.

INDUSTRIAL APPLICABILITY

[0103] According to the present invention, it is made possible to simply and rapidly detect cancer. In addition, the presence of a novel nucleic acid associated with the cancer is elucidated.

EQUIVALENT

[0104] Those skilled in the art will recognize, or be able to ascertain using simple routine experimentation, many equivalents to the specific embodiments of the invention described in the present specification. Such equivalents are intended to be encompassed in the scope of the following claims.

SEQUENCE LISTING

SEQ ID NO: 1
SEQUENCE LENGTH: 738
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

CCTCTGTCCA CTGCTTTCGT GAAGACAAGA TGAAGTTCAC AATTGTCTTT GCTGGACTTC 60
TTGGAGTCTT TCTAGCTCCT GCCCTTGCTA ACTATAATAT CAACGTCAAT GATGACAACA 120
ACAATGCTGG AAGTGGGCAG CAGTCAGTGA GTGTCAACAA TGAACACAAT GTGGCCAATG 180
TTGACAATAA CAACGGATGG GACTCCTGGA ATTCCATCTG GGATTATGGA AATGGCTTTG 240
CTGCAACCAG ACTCTTTCAA AAGAAGACAT GCATTGTGCA CAAAATGAAC AAGGAAGTCA 300
TGCCCTCCAT TCAATCCCTT GATGCACTGG TCAAGGAAAA GAAGCTTCAG GGTAAGGGAC 360
CAGGAGGACC ACCTCCCAAG GGCCTGATGT ACTCAGTCAA CCCAAACAAA GTCGATGACC 420
TGAGCAAGTT CGGAAAAAAC ATTGCAAACA TGTGTCGTGG GATTCCAACA TACATGGCTG 480
AGGAGATGCA AGAGGCAAGC CTGTTTTTTT ACTCAGGAAC GTGCTACACG ACCAGTGTAC 540
TATGGATTGT GGACATTTCC TTCTGTGGAG ACACGGTGGA GAACTAAACA ATTTTTTAAA 600
GCCACTATGG ATTTAGTCGT CTGAATATGC TGTGCAGAAA AAATATGGGC TCCAGTGGTT 660
TTTACCATGT CATTCTGAAA TTTTTCTCTA CTAGTTATGT TTGATTTCTT TAAGTTTCAA 720
TAAAATCATT TAGCATTG 738

SEQ ID NO: 2
SEQUENCE LENGTH: 2042
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

CCGTGACAAC ACTCCTGTCA TATTGGAGTC CAAAACTTGA ATTCTGGGTT GAATTTTTTA 60
AAAATCAGGT ACCACTTGAT TTCATATGGG AAATTGAAGC AGGAAATATT GAGGGCTTCT 120
TGATCACAGA AAACTCAGAA GAGATAGTAA TGCTCAGGAC AGGAGCGGCA GCCCCAGAAC 180
AGGCCACTCA TTTAGAATTC TAGTGTTTCA AAACACTTTT GTGTGTTGTA TGGTCAATAA 240

CATTTTTCAT TACTGATGGT GTCATTCACC CATTAGGTAA ACATTCCCTT TTAAATGTTT 300
GTTTGTTTTT TGAGACAGGA TCTCACTCTG TTGCCAGGGC TGTAGTGCAG TGGTGTGATC 360
ATAGCTCACT GCAACCTCCA CCTCCCAGGC TCAAGCCTCC CGAATAGCTG GGACTACAGG 420
CGCACACCAC CATCCCCGGC TAATTTTTGT ATTTTTTGTA GAGACGGGGT TTTGCCATGT 480
TGCCAAGGCT GGTTTCAAAC TCCTGGACTC AAGAAATCCA CCCACCTCAG CCTCCCAAAG 540
TGCTAGGATT ACAGGCATGA GCCACTGCGC CCAGCCCTTA TAAATTTTTG TATAGACATT 600
CCTTTGGTTG GAAGAATATT TATAGGCAAT ACAGTCAAAG TTTCAAAATA GCATCACACA 660
AAACATGTTT ATAAATGAAC AGGATGTAAT GTACATAGAT GACATTAAGA AAATTTGTAT 720
GAAATAATTT AGTCATCATG AAATATTTAG TTGTCATATA AAAACCCACT GTTTGAGAAT 780
GATGCTACTC TGATCTAATG AATGTGAACG TGTAGATGTT TTGTGTGTAT TTTTTTAAAT 840
GAAAACTCAA AATAAGACAA GTAATTTGTT GATAAATATT TTTAAAGATA ACTCAGCATG 900
TTTGTAAAGC AGGATACATT TTACTAAAAG GTTCATTGGT TCCAATCACA GCTCATAGGT 960
AGAGCAAAGA AAGGGTGGAT GGATTGAAAA GATTAGCNTN TGTNTCGGTG GCAGGTTCCC 1020
ACNTCGCAAG CAATTGGAAA CAAAANTTTN GGGGAGTTTT ATTTTGCATT NGGGTGTGTT 1080
TTATGTTAAG CAAAACATAN TTTAGAANCA AATGAAAAAG GCAATTGAAA ATCCCAGNTA 1140
TTTCACCTAG ATGGNATAGC CACCNTGAGC AGAACTTNGT GATGNTTCAT TCTGNGGAAT 1200
TTTGTGCTTN CTACTGTATA GTGCATGTGG TGTAGGTTAC TCTAACTGGT TTTGTNGACG 1260
TAAACATTTA AAGTGTTATA TTTTTTATAA AAATGTTTAT TTTTAATGAT ATGAGAAAAA 1320
TTTTGTTAGG CCACAAAAAC ACTGCACTGT GAACATTTTA GAAAAGGTAT GTCAGACTGG 1380
GATTAATGAC AGCATGATTT TCAATGACTG TAAATTGCGA TAAGGAAATG TACTGATTGC 1440
CAATACACCC CACCCTCATT ACATCATCAG GACTTGAAGC CAAGGGTTAA CCCAGCAAGC 1500
TACAAAGAGG GTGTGTCACA CTGAAACTCA ATAGTTGAGT TTGGCTGTTG TTGCAGGAAA 1560
ATGATTATAA CTAAAAGCTC TCTGATAGTG CAGAGACTTA CCAGAAGACA CAAGGAATTG 1620
TACTGAAGAG CTATTACAAT CCAAATATTG CCGTTTCATA AATGTAATAA GTAATACTAA 1680
TTCACAGAGT ATTGTAAATG GTGGATGACA AAAGAAAATC TGCTCTGTGG AAAGAAAGAA 1740
CTGTCTCTAC CAGGGTCAAG AGCATGAACG CATCAATAGA AAGRACTCGG GGAAACATCC 1800
CATCAACAGG ACTACACACT TGTATATACA TTCTTGAGAA CACTGCAATG TGAAAATCAC 1860
GTTTGCTATT TATAAACTTG TCCTTAGATT AATGTGTCTG GACAGATTGT GGGAGTAAGT 1920
GATTCTTCTA AGAATTAGAT ACTTGTCACT GCCTATACCT GCAGCTGAAC TGAATGGTAC 1980
TTCGKATGTT AATAGTTGTT CTGATAAATC ATGCAATTAA AATAAAGTGA TGCAACATCT 2040
TG 2042

SEQ ID NO: 3
SEQUENCE LENGTH: 1539

SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

ATGTTCGCCG ACCGTTGACT ATTCTCTACA AACCACAAAG ACATTGGAAC ACTATACCTA 60
TTATTCGGCG CATGAGCTGG AGTCCTAGGC ACAGCTCTAA GCCTCCTTAT TCGAGCCGAG 120
CTGGGCCAGC CAGGCAACCT TCTAGGTAAC GACCACATCT ACAACGTTAT CGTCACAGCC 180
CATGCATTTG TAATAATCTT CTTCATAGTA ATACCCATCA TAATCGGAGG CTTTGGCAAC 240
TGACTAGTTC CCCTAATAAT CGGTGCCCCC GATATGGCGT TCCCCCGCAT AAACAACATA 300
AGCTTCTGAC TCTTACCTCC CTCTCTCCTA CTCCTGCTCG CATCTGCTAT AGTAGAGGCC 360
GGAGCAGGAA CAGGTTGAAC AGTCTACCCT CCCTTAGCAG GGAACTACTC CCACCCTGGA 420
GCCTCCGTAG ACCTAACCAT CTTCTCCTTA CACCTAGCAG GTGTCTCCTC TATCTTAGGG 480
GCCATCAATT TCATCACAAC AATTATCAAT ATAAAACCCC CTGCCATAAC CCAATACCAA 540
ACGCCCCTCT TCGTCTGATC CGTCCTAATC ACAGCAGTCC TACTTCTCCT ATCTCTCCCA 600
GTCCTAGCTG CTGGCATCAC TATACTACTA ACAGACCGCA ACCTCAACAC CACCTTCTTC 660
GACCCCGCCG GAGGAGGAGA CCCCATTCTA TACCAACACC TATCCTGATT TTTCGGTCAC 720
CCTGAAGTTT ATATTCTTAT CCTACCAGGC TTCGGAATAA TCTCCCATAT TGTAACTTAC 780
TACTCCGGAA AAAAAGAACC ATTTGGATAC ATAGGTATGG TCTGAGCTAT GATATCAATT 840
GGCTTCCTAG GGTTTATCGT GTGAGCACAC CATATATTTA CAGTAGGAAT AGACGTAGAC 900
ACACGAGCAT ATTTCACCTC CGCTACCATA ATCATCGCTA TCCCCACCGG CGTCAAAGTA 960
TTTAGCTGAC TCGCCACACT CCACGGAAGC AATATGAAAT GATCTGCTGC AGTGCTCTGA 1020
GCCCTAGGAT TCATCTTTCT TTTCACCGTA GGTGGCCTGA CTGGCATTGT ATTAGCAAAC 1080
TCATCACTAG ACATCGTACT ACACGACACG TACTACGTTG TAGCTCACTT CCACTATGTC 1140
CTATCAATAG GAGCTGTATT TGCCATCATA GGAGGCTTCA TTCACTGATT TCCCCTATTC 1200
TCAGGCTACA CCCTAGACCA AACCTACGCC AAAATCCATT TCGCTATCAT ATTCATCGGC 1260
GTAAATCTAA CTTTCTTCCC ACAACACTTT CTCGGCCTAT CCGGAATGCC CCGACGTTAC 1320
TCGGACTACC CCGATGCATA CACCACATGA AATATCCTAT CATCTGTAGG CTCATTCATT 1380
TCTCTAACAG CAGTAATATT AATAATTTTC ATGATTTGAG AAGCCTTCGC TTCGAAGCGA 1440
AAAGTCCTAA TAGTAGAAGA ACCCTCCATA AACCTGGAGT GACTATATGG ATGCCCCCCA 1500
CCCTACCACA CATTCGAAGA ACCCGTATAC ATAAAATCT 1539

SEQ ID NO: 4
SEQUENCE LENGTH: 1807

SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

GAATTCTTTC TTCAGCCCAT GTAAACATGA AAATAAGGGT TAAAAATGAC TTCATTATGG 60
GGAAAAGGGA CAGGATGCAA ATTGTTCAAA TTCCGGGTGG CCGCTGCTCC GGCCTCCGGG 120
GCCTTGCGGA GACTCACCCC TTCAGCGTCG CTGCCCCCAG CTCAGCTCTT ACTGCGGGCC 180
GTCCGACGGC GGTCCCATCC TGTCAGGGAC TATGCGGCGC AAACATCTCC TTCGCCAAAA 240
GCAGGCGCCG CCACCGGGCG CATCGTGGCG GTCATTGGCG CAGTGGTGGA CGTCCAGTTT 300
GATGAGGGAC TACCACCAAT TCTAAATGCC CTGGAAGTGC AAGGCAGGGA GACCAGACTG 360
GTTTTGGAGG TGGCCCAGCA TTTGGGTGAG AGCACAGTAA GGACTATTGC TATGGATGGT 420
ACAGAAGGCT TGGTTAGAGG CCAGAAAGTA CTGGATTCTG GTGCACCAAT CAAAATTCCT 480
GTTGGTCCTG AGACTTTGGG CAGAATCATG AATGTCATTG GAGAACCTAT TGATGAAAGA 540
GGTCCCATCA AAACCAAACA ATTTGCTCCC ATTCATGCTG AGGCTCCAGA GTTCATGGAA 600
ATGAGTGTTG AGCAGGAAAT TCTGGTGACT GGTATCAAGG TTGTCGATCT GCTAGCTCCC 660
TATGCCAAGG GTGGCAAAAT TGGGCTTTTT GGTGGTGCTG GAGTTGGCAA GACTGTACTG 720
ATCATGGAGT TAATCAACAA TGTCGCCAAA GCCCATGGTG GTTACTCTGT GTTTGCTGGT 780
GTTGGTGAGA GGACCCGTGA AGGCAATGAT TTATACCATG AAATGATTGA ATCTGGTGTT 840
ATCAACTTAA AAGATGCCAC CTCTAAGGTA GCGCTGGTAT ATGGTCAAAT GAATCAACCA 900
CCTGGTGCTC GTGCCCGGGT AGCTCTGACT GGGCTGACTG TGGCTGAATA CTTCAGAGAC 960
CAAGAAGGTC AAGATGTACT GCTATTTATT GATAACATCT TTCGCTTCAC CCAGGCTGGT 1020
TCAGAGGTGT CTGCATTATT GGGCCGAATC CCTTCTGCTG TGGGCTATCA GCCTACCCTG 1080
GCCACTGACA TGGGCACTAT GCAGGAAAGA ATTACCACTA CCAAGAAGGG ATCTATCACC 1140
TCTGTACAGG CTATCTATGT GCCTGCTGAT GACTTGACTG ACCCTGCCCC TGCTACTACG 1200
TTTGCCCATT TGGATGCTAC CACTGTACTG TCGCGTGCCA TTGCTGAGCT GGGCATCTAT 1260
CCAGCTGTGG ATCCTCTAGA CTCCACCTCT CGTATCATGG ATCCCAACAT TGTTGGCAGT 1320
GAGCATTACG ATGTTGCCCG TGGGGTGCAA AAGATCCTGC AGGACTACAA ATCCCTCCAG 1380
GATATCATTG CCATCCTGGG TATGGATGAA CTTTCTGAGG AAGACAAGTT GACCGTGTCC 1440
CGTGCACGGA AAATACAGCG TTTCTTGTCT CAGCCATTCC AGGTTGCTGA GGTCTTCACA 1500
GGTCATATGG GGAAGCTGGT ACCCCTGAAG GAGACCATCA AAGGATTCCA GCAGATTTTG 1560
GCAGGTGAAT ATGACCATCT CCCAGAACAG GCCTTCTATA TGGTGGGACC CATTGAAGAA 1620
GCTGTGGCAA AAGCTGATAA GCTGGCTGAA GAGCATTCAT CGTGAGGGGT CTTTGTCCTC 1680
TGTACTTGTC TCTCTCCTTG CCCCTAACCC AAAAAGCTTC ATTTTTCTAT ATAGGCTGCA 1740

CAAGAGCCTT GATTGAAGAT ATATTCTTTC TGAACAGTAT TTAAGGTTTC CAATAAAATC 1800
GGAATTC 1807

SEQ ID NO: 5
SEQUENCE LENGTH: 4992
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

CCGCGGTGAG CCGCGAGGAA GAGAGGCGAG CGAGAGTGGA GGAGGAGGCG GCGGCTGCGG 60
GACGGTCCCC AGGAATGTCG CTGCCCCCCC CCCCCCTGCC GTTGAGGAGG AGACGGAGGA 120
GACCGACGTT GTTAGGGAAG ATGATCCCTA TGATCTGCCG CTGTTTCTGC ACAGAAATGA 180
GGGAAATACA AAGAACCAAA TACAGTTCTA AATTTGGGAT CTGTATTTTG AGATGATTTT 240
ATTTTCAGAA TGAGAAGCAT ATCTGGTTAC CTTTATGAAT GTAGAGACAT GAGAAGAGAG 300
TTATGATGGC AAAAAACAAA GAGCCTCGTC CCCCATCCTA TACCATCAGT ATAGTTGGAC 360
TCTCTGGGAC TGAAAAAGAC AAAGGTAACT GTGGAGTTGG AAAGTCTTGT TTGTGCAATA 420
GATTTGTACG CTCAAAAGCA GATGAATATT ATCCAGAGCA TACTTCTGTG CTTAGCACCA 480
TTGACTTTGG AGGACGAGTA GTAAACAATG ATCACTTTTT GTACTGGGGT GACATAATAC 540
AAAATAGTGA AGATGGAGTA GAATGCAAAA TTCATGTCAT TGAACAAACA GAGTTCATTG 600
ATGACCAGAC TTTCTTGCCT CATCGGAGTA CGAATTTGCA ACCATATATA AAACGTGCAG 660
CTGCATCTAA ATTGCAGTCA GCAGAAAAAC TAATGTACAT TTGCACTGAT CAGCTAGGCT 720
TAGAACAAGA CTTTGAACAG AAGCAAATGC CTGAAGGGAA GCTCAACGTA GATGGATTTT 780
TATTATGCAT TGATGTAAGT CAAGGATGCA ATAGGAAGTT TGATGATCAA CTTAAATTTG 840
TGAATAACCT TTTTGTCCAG TTATCAAAAT CAAAAAAACC TGTAATAATA GCAGCAACTA 900
AATGTGATGA ATGCGTGGGT CATTATCTTA GAGAAGTTCA GGCATTTGCT TCAAATAAAA 960
AGAACCTTCT TGTAGTGGAA ACACTCAGCG CAATAAAAGT CAACATTGAA ACATGTTTTA 1020
CTGCACTGGT ACAAATGTTG GATAAAACTC GTAGCAAGCC TAAAATTATT CCCTATTTGG 1080
ATGCTTATAA AACACAGAGA CAACTTGTTG TCACAGCAAC AGATAAGTTT GAAAAACTTG 1140
TGCAGACTGT GAGAGATTAT CATGCAACTT GGAAAACTGT TAGTAATAAA TTAAAAAATC 1200
ATCCTGATTA TGAAGAATAC ATCAACTTAG AGGGAACAAG AAAGGCCAGA AATACATTCT 1260
CAAAACATAT AGAACAACTT AAACAGGAAC ATATAAGAAA AAGGAGAGAA GAGTATATAA 1320
ATACTTTACC AAGAGCTTTT AACACTCTTT TGCCAAATCT AGAAGAGATT GAACATTTGA 1380
ATTGGTCAGA AGCTTTGAAG TTAATGGAAA AGAGAGCAGA TTTCCAGTTA TGTTTTGTGG 1440

TGCTAGAAAA AACTCCTTGG GATGAAACTG ACCATATAGA CAAAATTAAT GATAGGCGGA 1500
TTCCATTTGA CCTCCTGAGC ACTTTAGAAG CTGAAAAAGT CTATCAGAAC CATGTACAGC 1560
ATCTGATATC CGAGAAGAGG AGGGTGGAAA TGAAGGAAAA ATTCAAAAAG ACTTTGGAAA 1620
AAATTCAATT CATTTCACCA GGGCAGCCAT GGGAGGAAGT TATGTGCTTT GTTATGGAGG 1680
ATGAAGCCTA CAAATATATC ACTGAGGCTG ATAGCAAAGA GGTATATGGT AGGCATCAGC 1740
GAGAAATAGT TGAAAAAGCC AAAGAAGAGT TTCAAGAAAT GCTTTTTGAG CATTCTGAAC 1800
TTTTTTATGA TTTAGATCTT AATGCAACAC CTAGTTCAGA TAAAATGAGT GAAATTCATA 1860
CAGTTCTGAG TGAAGAACCT AGATATAAAG CTTTACAGAA ACTTGCACCT GATAGGGAAT 1920
CCCTTCTACT TAAGCATATA GGATTTGTTT ATCATCCCAC TAAAGAAACA TGTCTTAGTG 1980
GCCAAAATTG TACAGACATT AAAGTGGAGC AGTTACTTGC TAGTAGTCTT TTACAGTTGG 2040
ATCATGGCCG CTTAAGATTA TATCACGATA GTACCAATAT AGATAAAGTT AACCTTTTTA 2100
TTTTAGGGAA GGATGGCCTT GCCCAAGAAC TAGCAAATGA GATAAGGACA CAATCCACTG 2160
ATGATGAGTA TGCCTTAGAT GGAAAAATTT ATGAACTTGA TCTTCGGCCG GTTGATGCCA 2220
AATCGCCTTA CTTTTTGAGT CAGTTATGGA CTGCCGCCTT TAAACCACAT GGGTGCTTCT 2280
GTGTATTTAA TTCCATTGAG TCATTGAGTT TTATTGGGGA ATTTATTGGG AAAATAAGAA 2340
CTGAAGCTTC TCAGATCAGA AAAGATAAAT ACATGGCTAA TCTTCCATTT ACATTAATTC 2400
TGGCTAATCA GAGAGATTCC ATTAGTAAGA ATCTACCAAT TCTCAGGCAC CAAGGGCAGC 2460
AGTTGGCAAA CAAGTTGCAA TGTCCTTTTG TAGATGTACC TGCTGGTACA TATCCTCGTA 2520
AATTTAATGA AACCCAAATA AAGCAAGCTC TCAGAGGAGT ATTGGAATCA GTTAAACACA 2580
ATTTGGATGT GGTGAGCCCA ATTCCTGCCA ATAAGGACTT ATCAGAAGCT GACTTGAGAA 2640
TTGTCATGTG CGCCATGTGT GGAGATCCAT TTAGTGTGGA TCTTATTCTT TCACCCTTCC 2700
TTGATTCTCA TTCTTGCAGT GCTGCTCAAG CTGGACAGAA TAATTCCCTA ATGCTTGATA 2760
AAATCATTGG TGAAAAAAGG AGGCGAATAC AGATCACAAT ATTATCATAC CACTCTTCAA 2820
TTGGAGTAAG AAAAGATGAA CTAGTTCATG GGTATATATT AGTTTACTCT GCAAAACGGA 2880
AAGCTTCGAT GGGAATGCTT CGAGCATTTC TATCAGAAGT TCAAGACACC ATTCCTGTAC 2940
AGCTGGTGGC AGTTACTGAC AGCCAAGCAG ATTTTTTTGA AAATGAGGCT ATCAAAGAGT 3000
TAATGACTGA AGGAGAACAC ATTGCAACTG AGATCACTGC TAAATTTACA GCACTGTATT 3060
CTTTATCTCA GTATCATCGG CAAACTGAGG TCTTTACTCT GTTTTTTAGT GATGTTCTAG 3120
AGAAAAAAAA TATGATAGAA AATTCTTATT TGTCTGATAA TACAAGGGAA TCAACCCATC 3180
AAAGTGAAGA TGTTTTTCTA CCATCTCCCA GAGACTGTTT TCCCTATAAT AACTACCCTG 3240
ATTCAGATGA TGACACAGAA GCACCACCTC CTTATAGTCC AATTGGGGAT GATGTACAGT 3300
TGCTTCCAAC ACCTAGTGAC CGTTCCAGAT ATAGATTAGA TTTGGAAGGA AATGAGTATC 3360
CTATTCATAG TACCCCAAAC TGTCATGACC ATGAACGCAA CCATAAAGTG CCTCCACCTA 3420
TTAAACCTAA ACCAGTTGTA CCTAAGACAA ATGTGAAAGC GCTCGTTCCA AACCTTTTAA 3480

GGGCAATTGA AGCTGGTATT GGTAAAAATC CAAGAAAGCA GACTTCCCGG GTGCCTTTCG 3540
GTCCTGAAGA TATGGATCCT TCAGATAACT ATGCGGAACC CATTGATACA ATTTTCAAAC 3600
AGAAGGGCTA TTCTGATGAG ATTTATGTTG TCCCAGATGA TAGTCAAAAT CGTATTAAAA 3660
TTCGAAACTC ATTTGTAAAT AACACCCAAG GAGATGAAGA AAATGGGTTT TCTGATAGAC 3720
CTCAAAAAGT CATGGGGAAC GGAGGCCTTC AAAATACAAA TATAAATCTA AAACCTTGTT 3780
TAGTAAAGCC AAGTCATACT ATAGAAGAAC ACATTCAGAT GCCAGTGATG ATGAGGCTTT 3840
CACCACTTCT AAAACCAAAA AGAAAAGGAA GACATCGTGG AAGTGAAGAA GATCCACTTC 3900
TTTCTCCTGT TGAAACTTGG AAAGGTGGTA TTGATAATCC TGCAATCACT TCTGACCAGG 3960
AGTTAGATGA TAAGAAGATG AAGAAGAAAA CCCACAAAGT GAAAGAAGAT AAAAAAAAGA 4020
AAACTAAGAA CTTCAATCCA CCAACACGTA GAAATTGGGA AAGTAATTAC TTTGGGATGC 4080
CCCTCCAGGA TCTGGTTACA GCTGAGAAGC CCATACCACT ATTTGTTGAG AAATGTGTGG 4140
AATTTATTGA AGATACAGGG TTATGTACCG AGAGACTCTA CCGTGTCAGC GGGAATAAAA 4200
CTGACCAAGA AAATATTCAA AAGCAGTTTG TTCAAGATCA TAATATCAAT CTAGTGTCAA 4260
TGGAAGTAAC AGTAAATGCT GTAGCTGGAG CCCTTAAAGC TTTCTTTGCA GATCTGCCAG 4320
ATCCTTTAAT TCCATATTCT CTTCATCCAG AACTATTGGA AGCAGCAAAA ATCCCGGATA 4380
AAACAGAACG TCTTCATGCC TTGAAAGAAA TTGTTAAGAA ATTTCATCCT GTAAACTATG 4440
ATGTATTCAG ATACGTGATA ACACATCTAA ACAGGGTTAG TCAGCAACAT AAAATCAACC 4500
TAATGACAGC AGACAACTTA TCCATCTGTT TTGGCCAACC CTTGATGAGA CCTGATTTGA 4560
AATCGATGGA GTTTCTGTCT ACTACTAAGA TTCATCAATC TGTTGTTGAA ACATTCATTC 4620
AGCAGTGTCA GTTTTTCTTT TACAATGGAG AAATTGTAGA AACGACAAAC ATTGTGGCTC 4680
CTCCACCACC TTCAAACCCA GGACAGTTGG TGGAACCAAT GGTGCCACTT CAGTTGCCGC 4740
CACCATTGCA ACCTCAGCTG ATACAACCAC AATTACAAAC GGATCCTCTT GGTATTATAT 4800
GAGTAGGAAG TGATTGCAAA CAGGCTGGAT TTGGACAAAA AGCAAATCTA GACATGCATG 4860
TTTCAGGGTT CAGTAGTATA CTTCATGTTT CATACAGATA ATTCACATTC AAAATTACAT 4920
TTTCTCTTTG AACTAGATGG TATTCCTTAT TCACTTACAT TACAAATCTA AGACCATGTG 4980
ATAAGCATGA CT 4992

SEQ ID NO: 6
SEQUENCE LENGTH: 708
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

TATGGCACAT GCAGCGCAAG TAGGTCTACA AGACGCTACT TCCCCTATCA TAGAAGAGCT 60
TATCACCTTT CATGATCACG CCCTCATAAT CATTTTCCTT ATCTGCTTCC TAGTCCTGTA 120
TGCCCTTTTC CTAACACTCA CAACAAAACT AACTAATACT AACATCTCAG ACGCTCAGGA 180
AATAGAAACC GTCTGAACTA TCCTGCCCGC CATCATCCTA GTCCTCATCG CCCTCCCATC 240
CCTACGCATC CTTTACATAA CAGACGAGGT CAACGATCCC TCCCTTACCA TCAAATCAAT 300
TGGCCACCAA TGGTACTGAA CCTACGAGTA CACCGACTAC GGCGGACTAA TCTTCAACTC 360
CTACATACTT CCCCCATTAT TCCTAGAACC AGGCGACCTG CGACTCCTTG ACGTTGACAA 420
TCGAGTAGTA CTCCCGATTG AAGCCCCCAT TCGTATAATA ATTACATCAC AAGACGTCTT 480
GCACTCATGA GCTGTCCCCA CATTAGGCTT AAAAACAGAT GCAATTCCCG GACGTCTAAA 540
CCAAACCACT TTCACCGCTA CACGACCGGG GGTATACTAC GGTCAATGCT CTGAAATCTG 600
TGGAGCAAAC CACAGTTTCA TGCCCATCGT CCTAGAATTA ATTCCCCTAA AAATCTTTGA 660
AATAGGGCCC GTATTTACCC TATAGCACCC CCTCTACCCC CTCTAGAG 708

SEQ ID NO: 7
SEQUENCE LENGTH: 1140
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

ATGACCCCAA TACGCAAAAT TAACCCCCTA ATAAAATTAA TTAACCACTC ATTCATCGAC 60
CTCCCCACCC CATCCAACAT CTCCGCATGA TGAAACTTCG GCTCACTCCT TGGCGCCTGC 120
CTGATCCTCC AAATCACCAC AGGACTATTC CTAGCCATGC ACTACTCACC AGACGCCTCA 180
ACCGCCTTTT CATCAATCGC CCACATCACT CGAGACGTAA ATTATGGCTG AATCATCCGC 240
TACCTTCACG CCAATGGCGC CTCAATATTC TTTATCTGCC TCTTCCTACA CATCGGGCGA 300
GGCCTATATT ACGGATCATT TCTCTACTCA GAAACCTGAA ACATCGGCAT TATCCTCCTG 360
CTTGCAACTA TAGCAACAGC CTTCATAGGT TATGTCCTCC CGTGAGGCCA AATATCATTC 420
TGAGGGGCCA CAGTAATTAC AAACTTACTA TCCGCCATCC CATACATTGG GACAGACCTA 480
GTTCAATGAA TCTGAGGAGG CTACTCAGTA GACAGTCCCA CCCTCACACG ATTCTTTACC 540
TTTCACTTCA TCTTGCCCTT CATTATTGCA ACCCTAGCAG CACTCCACCT CCTATTCTTG 600
CACGAAACGG GATCAAACAA CCCCCTAGGA ATCACCTCCC ATTCCGATAA AATCACCTTC 660
CACCCTTACT ACACAATCAA AGACACCCTC GGCTTACTTC TCTTCCTTCT CTCCTTAATG 720
ACATTAACAC TATTCTCACC AGACCTCCTA GGCGACCCAG ACAATTATAC CCTAGCCAAC 780
CCCTTAAACA CCCCTCCCCA CATCAAGCCC GAATGATATT TCCTATTCGC CTACACAATT 840

CTCCGATCCG TCCCTAACAA ACTAGGAGGC GTCCTTGCCC TATTACTATC CATCCTCATC 900
CTAGCAATAA TCCCCATCCT CCATATATCC AAACAACAAA GCATAATATT TCGCCCACTA 960
AGCCAATCAC TTTATTGACT CCTAGCCGCA GACCTCCTCA TTCTAACCTG AATCGGAGGA 1020
CAACCAGTAA GCTACCCTTT TACCATCATT GGACAAGTAG CATCCGTACT ATACTTCACA 1080
ACAATCCTAA TCCTAATACC AACTATCTCC CTAATTGAAA ACAAAATACT CAAATGGGCC 1140

SEQ ID NO: 8
SEQUENCE LENGTH: 5629
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

GCGCGACCGT CCCGGGGGTG GGGCCGGGCG CAGCGGCGAG AGGAGGCGAA GGTGGCTGCG 60
GTAGCAGCAG CGCGGCAGCC TCGGACCCAG CCCGGAGCGC AGGGCGGCCG CTGCAGGTCC 120
CCGCTCCCCT CCCCGTGCGT CCGCCCATGG CCGCCGCCGG GCAGCTGTGC TTGCTCTACC 180
TGTCGGCGGG GCTCCTGTCC CGGCTCGGCG CAGCCTTCAA CTTGGACACT CGGGAGGACA 240
ACGTGATCCG GAAATATGGA GACCCCGGGA GCCTCTTCGG CTTCTCGCTG GCCATGCACT 300
GGCAACTGCA GCCCGAGGAC AAGCGGCTGT TGCTCGTGGG GGCCCCGCGC GGAGAAGCGC 360
TTCCACTGCA GAGAGCCAAC AGAACGGGAG GGCTGTACAG CTGCGACATC ACCGCCCGGG 420
GGCCATGCAC GCGGATCGAG TTTGATAACG ATGCTGACCC CACGTCAGAA AGCAAGGAAG 480
ATCAGTGGAT GGGGGTCACC GTCCAGAGCC AAGGTCCAGG GGGCAAGGTC GTGACATGTG 540
CTCACCGATA TGAAAAAAGG CAGCATGTTA ATACGAAGCA GGAATCCCGA GACATCTTTG 600
GGCGGTGTTA TGTCCTGAGT CAGAATCTCA GGATTGAAGA CGATATGGAT GGGGGAGATT 660
GGAGCTTTTG TGATGGGCGA TTGAGAGGCC ATGAGAAATT TGGCTCTTGC CAGCAAGGTG 720
TAGCAGCTAC TTTTACTAAA GACTTTCATT ACATTGTATT TGGAGCCCCG GGTACTTATA 780
ACTGGAAAGG GATTGTTCGT GTAGAGCAAA AGAATAACAC TTTTTTTGAC ATGAACATCT 840
TTGAAGATGG GCCTTATGAA GTTGGTGGAG AGACTGAGCA TGATGAAAGT CTCGTTCCTG 900
TTCCTGCTAA CAGTTACTTA GGTTTTTCTT TGGACTCAGG GAAAGGTATT GTTTCTAAAG 960
ATGAGATCAC TTTTGTATCT GGTGCTCCCA GAGCCAATCA CAGTGGAGCC GTGGTTTTGC 1020
TGAAGAGAGA CATGAAGTCT GCACATCTCC TCCCTGAGCA CATATTCGAT GGAGAAGGTC 1080
TGGCCTCTTC ATTTGGCTAT GATGTGGCGG TGGTGGACCT CAACAAGGAT GGGTGGCAAG 1140
ATATAGTTAT TGGAGCCCCA CAGTATTTTG ATAGAGATGG AGAAGTTGGA GGTGCAGTGT 1200
ATGTCTACAT GAACCAGCAA GGCAGATGGA ATAATGTGAA GCCAATTCGT CTTAATGGAA 1260

CCAAAGATTC TATGTTTGGC ATTGCAGTAA AAAATATTGG AGATATTAAT CAAGATGGCT 1320
ACCCAGATAT TGCAGTTGGA GCTCCGTATG ATGACTTGGG AAAGGTTTTT ATCTATCATG 1380
GATCTGCAAA TGGAATAAAT ACCAAACCAA CACAGGTTCT CAAGGGTATA TCACCTTATT 1440
TTGGATATTC AATTGCTGGA AACATGGACC TTGATCGAAA TTCCTACCCT GATGTTGCTG 1500
TTGGTTCCCT CTCAGATTCA GTAACTATTT TCAGATCCCG GCCTGTGATT AATATTCAGA 1560
AAACCATCAC AGTAACTCCT AACAGAATTG ACCTCCGCCA GAAAACAGCG TGTGGGGCGC 1620
CTAGTGGGAT ATGCCTCCAG GTTAAATCCT GTTTTGAATA TACTGCTAAC CCCGCTGGTT 1680
ATAATCCTTC AATATCAATT GTGGGCACAC TTGAAGCTGA AAAAGAAAGA AGAAAATCTG 1740
GGCTATCCTC AAGAGTTCAG TTTCGAAACC AAGGTTCTGA GCCCAAATAT ACTCAAGAAC 1800
TAACTCTGAA GAGGCAGAAA CAGAAAGTGT GCATGGAGGA AACCCTGTGG CTACAGGATA 1860
ATATCAGAGA TAAACTGCGT CCCATTCCCA TAACTGCCTC AGTGGAGATC CAAGAGCCAA 1920
GCTCTCGTAG GCGAGTGAAT TCACTTCCAG AAGTTCTTCC AATTCTGAAT TCAGATGAAC 1980
CCAAGACAGC TCATATTGAT GTTCACTTCT TAAAAGAGGG ATGTGGAGAC GACAATGTAT 2040
GTAACAGCAA CCTTAAACTA GAATATAAAT TTTGCACCCG AGAAGGAAAT CAAGACAAAT 2100
TTTCTTATTT ACCAATTCAA AAAGGTGTAC CAGAACTAGT TCTAAAAGAT CAGAAGGATA 2160
TTGCTTTAGA AATAACAGTG ACAAACAGCC CTTCCAACCC AAGGAATCCC ACAAAAGATG 2220
GCGATGACGC CCATGAGGCT AAACTGATTG CAACGTTTCC AGACACTTTA ACCTATTCTG 2280
CATATAGAGA ACTGAGGGCT TTCCCTGAGA AACAGTTGAG TTGTGTTGCC AACCAGAATG 2340
GCTCGCAAGC TGACTGTGAG CTCGGAAATC CTTTTAAAAG AAATTCAAAT GTCACTTTTT 2400
ATTTGGTTTT AAGTACAACT GAAGTCACCT TTGACACCCC ATATCTGGAT ATTAATCTGA 2460
AGTTAGAAAC AACAAGCAAT CAAGATAATT TGGCTCCAAT TACAGCTAAA GCAAAAGTGG 2520
TTATTGAACT GCTTTTATCG GTCTCGGGAG TTGCTAAACC TTCCCAGGTG TATTTTGGAG 2580
GTACAGTTGT TGGCGAGCAA GCTATGAAAT CTGAAGATGA AGTGGGAAGT TTAATAGAGT 2640
ATGAATTCAG GGTAATAAAC TTAGGTAAAC CTCTTACAAA CCTCGGCACA GCAACCTTGA 2700
ACATTCAGTG GCCAAAAGAA ATTAGCAATG GGAAATGGTT GCTTTATTTG GTGAAAGTAG 2760
AATCCAAAGG ATTGGAAAAG GTAACTTGTG AGCCACAAAA GGAGATAAAC TCCCTGAACC 2820
TAACGGAGTC TCACAACTCA AGAAAGAAAC GGGAAATTAC TGAAAAACAG ATAGATGATA 2880
ACAGAAAATT TTCTTTATTT GCTGAAAGAA AATACCAGAC TCTTAACTGT AGCGTGAACG 2940
TGAACTGTGT GAACATCAGA TGCCCGCTGC GGGGGCTGGA CAGCAAGGCG TCTCTTATTT 3000
TGCGCTCGAG GTTATGGAAC AGCACATTTC TAGAGGAATA TTCCAAACTG AACTACTTGG 3060
ACATTCTCAT GCGAGCCTTC ATTGATGTGA CTGCTGCTGC CGAAAATATC AGGCTGCCAA 3120
ATGCAGGCAC TCAGGTTCGA GTGACTGTGT TTCCCTCAAA GACTGTAGCT CAGTATTCGG 3180
GAGTACCTTG GTGGATCATC CTAGTGGCTA TTCTCGCTGG GATCTTGATG CTTGCTTTAT 3240
TAGTGTTTAT ACTATGGAAG TGTGGTTTCT TCAAGAGAAA TAAGAAAGAT CATTATGATG 3300

CCACATATCA CAAGGCTGAG ATCCATGCTC AGCCATCTGA TAAAGAGAGG CTTACTTCTG 3360
ATGCATAGTA TTGATCTACT TCTGTAATTG TGTGGATTCT TTAAACGCTC TAGGTACGAT 3420
GACAGTGTTC CCCGATACCA TGCTGTAAGG ATCCGGAAAG AAGAGCGAGA GATCAAAGAT 3480
GAAAAGTATA TTGATAACCT TGAAAAAAAA CAGTGGATCA CAAAGTGGAA CAGAAATGAA 3540
AGCTACTCAT AGCGGGGGCC TAAAAAAAAA AAAGCTTCAC AGTACCCAAA CTGCTTTTTC 3600
CAACTCAGAA ATTCAATTTG GATTTAAAAG CCTGCTCAAT CCCTGAGGAC TGATTTCAGA 3660
GTGACTACAC ACAGTACGAA CCTACAGTTT TAACTGTGGA TATTGTTACG TAGCCTAAGG 3720
CTCCTGTTTT GCACAGCCAA ATTTAAAACT GTTGGAATGG ATTTTTCTTT AACTGCCGTA 3780
ATTTAACTTT CTGGGTTGCC TTTGTTTTTG GCGTGGCTGA CTTACATCAT GTGTTGGGGA 3840
AGGGCCTGCC CAGTTGCACT CAGGTGACAT CCTCCAGATA GTGTAGCTGA GGAGGCACCT 3900
ACACTCACCT GCACTAACAG AGTGGCCGTC CTAACCTCGG GCCTGCTGCG CAGACGTCCA 3960
TCACGTTAGC TGTCCCACAT CACAAGACTA TGCCATTGGG GTAGTTGTGT TTCAACGGAA 4020
AGTGCTGTCT TAAACTAAAT GTGCAATAGA AGGTGATGTT GCCATCCTAC CGTCTTTTCC 4080
TGTTTCCTAG CTGTGTGAAT ACCTGCTCAC GTCAAATGCA TACAAGTTTC ATTCTCCCTT 4140
TCACTAAAAA CACACAGGTG CAACAGACTT GAATGCTAGT TATACTTATT TGTATATGGT 4200
ATTTATTTTT TCTTTTCTTT ACAAACCATT TTGTTATTGA CTAACAGGCC AAAGAGTCTC 4260
CAGTTTACCC TTCAGGTTGG TTTAATCAAT CAGAATTAGA ATTAGAGCAT GGGAGGGTCA 4320
TCACTATGAC CTAAATTATT TACTGCAAAA AGAAAATCTT TATAAATGTA CCAGAGAGAG 4380
TTGTTTTAAT AACTTATCTA TAAACTATAA CCTCTCCTTC ATGACAGCCT CCACCCCACA 4440
ACCCAAAAGG TTTAAGAAAT AGAATTATAA CTGTAAAGAT GTTTATTTCA GGCATTGGAT 4500
ATTTTTTACT TTAGAAGCCT GCATAATGTT TCTGGATTTA CATACTGTAA CATTCAGGAA 4560
TTCTTGGAGA AGATGGGTTT ATTCACTGAA CTCTAGTGCG GTTTACTCAC TGCTGCAAAT 4620
ACTGTATATT CAGGACTTGA AAGAAATGGT GAATGCCTAT GGAACTAGTG GATCCAAACT 4680
GATCCAGTAT AAGACTACTG AATCTGCTAC CAAAACAGTT AATCAGTGAG TCGAGTGTTC 4740
TATTTTTTGT TTTGTTTCCT CCCCTATCTG TATTCCCAAA AATTACTTTG GGGCTAATTT 4800
AACAAGAACT TTAAATTGTG TTTTAATTGT AAAAATGGCA GGGGGTGGAA TTATTACTCT 4860
ATACATTCAA CAGAGACTGA ATAGATATGA AAGCTGATTT TTTTTAATTA CCATGCTTCA 4920
CAATGTTAAG TTATATGGGG AGCAACAGCA AACAGGTGCT AATTTGTTTT GGATATAGTA 4980
TAAGCAGTGT CTGTGTTTTG AAAGAATAGA ACACAGTTTG TAGTGCCACT GTTGTTTTGG 5040
GGGGGGCTTT TTTTCTTTTT CCGGAAAATC CTTAAACCTT AAGATACTAA GGACGTTGTT 5100
TTGGTTGTAC TTGGAATTCT TAGTCACAAA ATATATTTTG TTTACAAAAA TTTCTGTAAA 5160
ACAGGTTATA ACAGTGTTTA AAGTCTCAGT TTCTTGCTTG GGGAACTTGT GTCCCTAATG 5220
TGTTAGATTG CTAGATTGCT AAGGAGCTGA TACTTGACAG TTTTTTAGAC CTGTGTTACT 5280
AAAAAAAAGA TGAATGTCGG AAAAGGGTGT TGGGAGGGTG GTCAACAAAG AAACAAAGAT 5340

GTTATGGTGT TTAGACTTAT GGTTGTTAAA AATGTCATCT CAAGTCAAGT CACTGGTCTG 5400
TTTGCATTTG ATACATTTTT GTACTAACTA GCATTGTAAA ATTATTTCAT GATTAGAAAT 5460
TACCTGTGGA TATTTGTATA AAAGTGTGAA ATAAATTTTT TATAAAAGTG TTCATTGTTT 5520
CGTAACACAG CATTGTATAT GTGAAGCAAA CTCTAAAATT ATAAATGACA ACCTGAATTA 5580
TCTATTTCAT CAAAAAAAAA AAAAAAAAAA ACTTTATGGG CACAACTGG 5629

SEQ ID NO: 9
SEQUENCE LENGTH: 580
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:
CCATCCAATG AGGCCACCTC TTTCTAAACT CAGACTCTTC ATTTAGGGAG GTGAGTTCCA 60
TTAAGGAACT TGAGATTTTC AGATAAATGG AAAATACTAG ATAAAGAGGT ATCTCATAGA 120
TAGCAAAGGT AAACTCTCAT ACAATCATTG AGCTAGGACA TTAATGGTTC AGTGGTTCCC 180
AATTCTAGAT ATACATTAAA ATAAATTGAA AAGCCTTTTA AAAATACATG ATTACTGGAC 240
CTACTGAATT ATATCCTTTG GGGAGCCCAA GAACTTATTA AATTCTCTGG GCTATTTTTA 300
TGATTTCTCT GAGCTGTTAC TGGGAACTAC TGATTGAATC CATYTTTTAT AGTAATGTTT 360
CCAACAGAAG GCTGTTTSCC TTTGCTTAAC ATTATTTCCA GTGAAGTATT ATTTTCCATT 420
CTGGAGACAG TTCAAAAGTT TTTTTAAGTA ACAGCTTTAT TGAGACAATT TATATSCCGT 480
ACAATTCACC TAAAGTGTGT AATTCAGTTG TTTTTAGTAT GTTCACAGAA TTGTGCAGCT 540
TGCATCTATC ACCACAAATT TAGAACCTTG TCATAATCCC 580

SEQ ID NO: 10
SEQUENCE LENGTH: 1552
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:
CCCAAACCCA CTCCACCTTA CTACCAGACA ACCTTAGCCA AACCATTTAC CCAAATAAAG 60
TATAGGCGAT AGAAATTGAA ACCTGGCGCA ATAGATATAG TACCGCAAGG GAAAGATGAA 120
AAATTATAAC CAAGCATAAT ATAGCAAGGA CTAACCCCTA TACCTTCTGC ATAATGAATT 180

AACTAGAAAT AACTTTGCAA GGAGAGCCAA AGCTAAGACC CCCGAAACCA GACGAGCTAC 240
CTAAGAACAG CTAAAAGAGC ACACCCGTCT ATGTAGCAAA ATAGTGGGAA GATTTATAGG 300
TAGAGGCGAC AAACCTACCG AGCCTGGTGA TAGCTGGTTG TCCAAGATAG AATCTTAGTT 360
CAACTTTAAA TTTGCCCACA GAACCCTCTA AATCCCCTTG TAAATTTAAC TGTTAGTCCA 420
AAGAGGAACA GCTCTTTGGA CACTAGGAAA AAACCTTGTA GAGAGAGTAA AAAATTTAAC 480
ACCCATAGTA GGCCTAAAAG CAGCCACCAA TTAAGAAAGC GTTCAAGCTC AACACCCACT 540
ACCTAAAAAA TCCCAAACAT ATAACTGAAC TCCTCACACC CAATTGGACC AATCTATCAC 600
CCTATAGAAG AACTAATGTT AGTATAAGTA ACATGAAAAC ATTCTCCTCC GCATAAGCCT 660
GCGTCAGATT AAAACACTGA ACTGACAATT AACAGCCCAA TATCTACAAT CAACCAACAA 720
GTCATTATTA CCCTCACTGT CAACCCAACA CAGGCATGCT CATAAGGAAA GGTTAAAAAA 780
AGTAAAAGGA ACTCGGCAAA TCTTACCCCG CCTGTTTACC AAAAACATCA CCTCTAGCAT 840
CACCAGTATT AGAGGCACCG CCTGCCCAGT GACACATGTT TAACGGCCGC GGTACCCTAA 900
CCGTGCAAAG GTAGCATAAT CACTTGTTCC TTAATTAGGG ACCCGTATGA ATGGCTCCAC 960
GAGGGTTCAG CTGTCTCTTA CTTTTAACCA GTGAAATTGA CCTGCCCGTG AAGAGGCGGG 1020
CATGACACAG CAAGACGAGA AGACCCTATG GAGCTTTAAT TTATTAATGC AAACAGTACC 1080
TAACAAACCT ACAGGTCCTA AACTACCAAA CCTGCATTAA AAATTTCGGT TGGGGCGACC 1140
TCGGAGCAGA ACCCAACCTC CGAGCAGTAC ATGCTAAGAC TTCACCAGTC AAAGCGAACT 1200
ACTATACTCA ATTGATCCAA TAACTTGACC AACGGAACAA GTTACCCTAG GGATAACAGC 1260
GCAATCCTAT TCTAGAGTCC ATATCAACAA TAGGGTTTAC GACCTCGATG TTGGATCAGG 1320
ACATCCCGAT GGTGCAGCCG CTATTAAAGG TTCGTTTGTT CAACGATTAA AGTCCTACGT 1380
GATCTGAGTT CAGACCGGAG TAATCCAGGT CGGTTTCTAT CTACTTCAAA TTCCTCCCTG 1440
TACGAAAGGA CAAGAGAAAT AAGGCCTACT TCACAAAGCG CCTTCCCCCG TAAATGATAT 1500
CATCTCAACT TAGTATTATA CCCACACCCA CCCAAGAACA GGGTTTGTTA AG 1552

SEQ ID NO: 11
SEQUENCE LENGTH: 2116
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:
GGGTGGCAGA ATATTAGTCT AGCTATCTCC CATTGCTCTC ACGCGCCATC TACTGGATTT 60
CATCCCAAAC TACAACACGA AAAACTGCTA ATTTTCCTGC CTGCCAGGCC GAGGACTGGA 120
ATTCAACAGA CTGTTTAGAG CCTTTGCCCT CTGAAAACTT CCAGAAATGA AGCCAACTGA 180

CTATATTCAG TTTACACCAG AGTTAAAGGA ACGCCAACCC TCCCAGATGA GAAAGAATCA 240
GTGCAAGAAC TGTAGCAATT TAAAAAACCA GAGCGTCCCC TTACCTCCAA ATGAGCCCAC 300
TAGCTCCACA GCAATTGTTC TTAACCAATC TGAAATGATG AGCATGGAAT TCAGAATCTG 360
AATGGCAATG AAGCTTATAG ATATCCAGGA GAAAGTTGAA ATGCAATCCA AGGAAACCAA 420
GCAATCCAGT GAAATGGTTT AAGAGCTGAA AGATAAAATA NCAATTTTAC AAAAGACCCA 480
AACTGAGCTT ATTGAGTTCA AAAAAGAATT TCATAATACA ATCAGAAGTA TTAATAGCAG 540
AATAGGCCAA GCTGAGGAAA GAATCTCAGA GCTTGACCCC TGGTTCTTTG AATCAACTTA 600
GACAAAAATA AAGAAAAAAG AGTTTTAAGA AATGAACACA ATCTCCCAGA AATATGAGAT 660
TATGTWAAGA GACAAAATCT ATGACTCATT GCCATCCCTG AGAGAGAAGG AGAGAGAATA 720
AGCAACTTGG AAAATATATT TGGGGACATA GCCCACAAAA ATTTCCCTAA TCTCTCTAGA 780
GAGGTTGACA TGTAAATTCA AGAAATACAG AAGACCTTGG CCAGATAATA TACAAGATGA 840
CCATCCCCAA GGCACATAGT CATCAGATTC ACCATGGTCA ATGCAAAAGA AAAAAATCTT 900
AAAGACAGCT AGGGAGAAGG GTCAAGTCAC ATGCAGAAGG ACTCTCATTA GGCTGGCAGT 960
GGACCTCTCA GCAGAAACCT GACAAGCCAG AAGAGATGGA GGGAGAGGGG TCTATTTTTG 1020
TCATCCTTAA AGAAAAAAAA TTCCAACCAA GAGTCTCATA CACTGCCAAA CTAAGCTTCC 1080
TAAGTGAAGG AGAAATAAAA ACCTTCTCAG ACAAGCAAAT GCTGAAGGAA TTCAACTAGA 1140
CCAGCCTAAC AAGAGGTCCT AAGGGAGTGC TGAATATGGA CTCAAAAGAA TAACACCTGC 1200
TACCACAAAC ACTCACTTAA GCACACAGCC CAACGACACT ATAGGCAATT ACACAGTAAG 1260
TCTACATAAC AACACAATGA CAGGATCAAC ATCTCACACA TCAATACTAA CCCCGAGTGT 1320
AAAGGGGCTA AATGCCCCAC TTAAAAGACA TAGAGTGTCA AGCTTGATAA AAAGACAAGA 1380
TCCAATCATC CACTATTTTC AAGAGCTCTA TGTTATGTGT AATGACACCC ACAGACTCAA 1440
AGACTTGGAG AAAGATTTAT CATGCAAAAT CAGAAAACAA AAAAGAGCAG GAGTCACTAG 1500
TTTTATATCA GACAAAACAG ACTTTAAACC CTTAATAATT AAGAAAGACA AAGAAGGGTA 1560
TTTCCTGGAC CACAGAAGGC TTATTGGAAA AAAGGACATA ATGACAAAGG GTACAATCCA 1620
ACAAGAAGTT TTAACTATTC TAAATATATA CACACCCAAC ATTGGAGCAC CCAGATTTAT 1680
AAAACAAGTA CTTCTCGATC TACAAGAAGA CTTAGACAGC CACACAATAA TAGTGGGAGA 1740
CTTTCACATC CTACTTACAG ATCATTGAGA CAGAAAACTA ATAAAAGAAC TCTGGACTTA 1800
AACTTGTTAC TTGACCAATT GGACCTAATA GATATCCACA GAAAACTTCA CCCAACAAAG 1860
ACAGAATATA CATTCTTCTT ATCTGCACAT GGAACACATT CCAAGATCAA TCACATGCTA 1920
GGTAAGAAAG CAAGTCTCAA TAAATTAAAA AAAATTGAAA TCATACGAAC CTTAATATCA 1980
GACCACAATG TAATTAAAAA TAAATCAATA TCAAGAAGAT CTCATACATA AATACATGAA 2040
AATTAAACAA CTTACTCCTG AATAACTCTT GTGTGAACAT CAAAATTCAG GAAGAAATAA 2100
AAAATTATTT GAAATT 2116

SEQ ID NO: 12
SEQUENCE LENGTH: 173
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

GCGATCCACA AATGGGAGGT GACGGTCCAT CAGGGAAGCT GGGTTCGCGG CTCCACGGCT 60
GGGGGCTGCC GCAATTTCCT GGATACCTTT TGGACCAATC CACAAATAAA ATTGTCTCTG 120
ACTGAGAAAG ATGAGGGGCA GGAGGAGTGT AGTTTCCTTG TAGCCCTGAT GCA 173

SEQ ID NO: 13
SEQUENCE LENGTH: 655
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

CTGATCCATG GGCCAGCAGC ATCAATATTA CCTGGGAGCT TACAGAAATG CAGAATTTCA 60
GGCCCACTGC AGATCTACCG AATCAAAATC TTCCTTTAGC AAAATTTCTC AAACGATTAG 120
CACTGGCCTA CATCCATTTT ATCCTTCCTT AGCTATTAGG GATGTGAGGT CCGAGGGCTT 180
CAAAAGGTCC CCGGAATAGC TTGTTCCTTC ATCCACTGTG TCCTATTCAT TCTTCAGCTA 240
ACTCCAGCAA TGAGCTGAAA CTCATTCATC ACCCTTGCTG AGTTTTCTTC TCAATCCTTA 300
TTCCTAATTC TGGTTCTAGA TGAGCCCTAC CTACCCAGTG GTTGTATTTT TGTAGCCAGT 360
GTGGGACACA GGAGATTGGC AGACCAACAC AGCTAGCCTC TCTCTAGCCC TCCCTCCACC 420
TCTAAGTCAC TAACAATCCA TGTTTGTTCA GTTTGTTGAC ATGTGGCATG TTCATTTGTT 480
CACAACTTAA TCACGGGGGA CATTTCAGAA AAATGTGTAC TAAGTTAAAA CCATGTTTAG 540
TCTCCTACAA CTTGTACATT TTCATTTTCT CTTATCAGTA GATTGTCCTT GTTGACATAG 600
CTCATGCATG AGGACACATA GCAGTACACA CACATTGAAT GAATTGTTAG TCATG 655

SEQ ID NO: 14
SEQUENCE LENGTH: 2619
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear

MOLECULAR TYPE: cDNA to mRNA

SEQUENCE DESCRIPTION:

GACTCCTAGG GGCTTGCAGA CCTAGTGGGA GAGAAAGAAC ATCGCAGCAG CCAGGCAGAA 60
CCAGGACAGG TGAGGTGCAG GCTGGCTTTC CTCTCGCAGC GCGGTGTGGA GTCCTGTCCT 120
GCCTCAGGGC TTTTCGGAGC CTGGATCCTC AAGGAACAAG TAGACCTGGC CGCGGGGAGT 180
GGGGAGGGAA GGGGTGTCTA TTGGGCAACA GGGCGGCAAA GCCCTGAATA AAGGGGCGCA 240
GGGCAGGCGC AAGTGCAGAG CCTTCGTTTG CCAAGTCGCC TCCAGACCGC AGACATGAAA 300
CTTGTCTTCC TCGTCCTGCT GTTCCTCGGG GCCCTCGGAC TGTGTCTGGC TGGCCGTAGG 360
AGAAGGAGTG TTCAGTGGTG CGCCGTATCC CAACCCGAGG CCACAAAATG CTTCCAATGG 420
CAAAGGAATA TGAGAAAAGT GCGTGGCCCT CCTGTCAGCT GCATAAAGAG AGACTCCCCC 480
ATCCAGTGTA TCCAGGCCAT TGCGGAAAAC AGGGCCGATG CTGTGACCCT TGATGGTGGT 540
TTCATATACG AGGCAGGCCT GGCCCCCTAC AAACTGCGAC CTGTAGCGGC GGAAGTCTAC 600
GGGACCGAAA GACAGCCACG AACTCACTAT TATGCCGTGG CTGTGGTGAA GAAGGGCGGC 660
AGCTTTCAGC TGAACGAACT GCAAGGTCTG AAGTCCTGCC ACACAGGCCT TCGCAGGACC 720
GCTGGATGGA ATGTCCCTAC AGGGACACTT CGTCCATTCT TGAATTGGAC GGGTCCACCT 780
GAGCCCATTG AGGCAGCTGT GGCCAGGTTC TTCTCAGCCA GCTGTGTTCC CGGTGCAGAT 840
AAAGGACAGT TCCCCAACCT GTGTCGCCTG TGTGCGGGGA CAGGGGAAAA CAAATGTGCC 900
TTCTCCTCCC AGGAACCGTA CTTCAGCTAC TCTGGTGCCT TCAAGTGTCT GAGAGACGGG 960
GCTGGAGACG TGGCTTTTAT CAGAGAGAGC ACAGTGTTTG AGGACCTGTC AGACGAGGCT 1020
GAAAGGGACG AGTATGAGTT ACTCTGCCCA GACAACACTC GGAAGCCAGT GGACAAGTTC 1080
AAAGACTGCC ATCTGGCCCG GGTCCCTTCT CATGCCGTTG TGGCACGAAG TGTGAATGGC 1140
AAGGAGGATG CCATCTGGAA TCTTCTCCGC CAGGCACAGG AAAAGTTTGG AAAGGACAAG 1200
TCACCGAAAT TCCAGCTCTT TGGCTCCCCT AGTGGGCAGA AAGATCTGCT GTTCAAGGAC 1260
TCTGCCATTG GGTTTTCGAG GGTGCCCCCG AGGATAGATT CTGGGCTGTA CCTTGGCTCC 1320
GGCTACTTCA CTGCCATCCA GAACTTGAGG AAAAGTGAGG AGGAAGTGGC TGCCCGGCGT 1380
GCGCGGGTCG TGTGGTGTGC GGTGGGCGAG CAGGAGCTGC GCAAGTGTAA CCAGTGGAGT 1440
GGCTTGAGCG AAGGCAGCGT GACCTGCTCC TCGGCCTCCA CCACAGAGGA CTGCATCGCC 1500
CTGGTGCTGA AAGGAGAAGC TGATGCCATG AGTTTGGATG GAGGATATGT GTACACTGCA 1560
TGCAAATGTG GTTTGGTGCC TGTCCTGGCA GAGAACTACA AATCCCAACA AAGCAGTGAC 1620
CCTGATCCTA ACTGTGTGGA TAGACCTGTG GAAGGATATC TTGCTGTGGC GGTGGTTAGG 1680
AGATCAGACA CTAGCCTTAC CTGGAACTCT GTGAAAGGCA AGAAGTCCTG CCACACCGCC 1740
GTGGACAGGA CTGCAGGCTG GAATATCCCC ATGGGCCTGC TCTTCAACCA GACGGGCTCC 1800
TGCAAATTTG ATGAATATTT CAGTCAAAGC TGTGCCCCTG GGTCTGACCC GAGATCTAAT 1860

CTCTGTGCTC TGTGTATTGG CGACGAGCAG GGTGAGAATA AGTGCGTGCC CAACAGCAAC 1920
GAGAGATACT ACGGCTACAC TGGGGCTTTC CGGTGCCTGG CTGAGAATGC TGGAGACGTT 1980
GCATTTGTGA AAGATGTCAC TGTCTTGCAG AACACTGATG GAAATAACAA TGAGGCATGG 2040
GCTAAGGATT TGAAGCTGGC AGACTTTGCG CTGCTGTGCC TCGATGGCAA ACGGAAGCCT 2100
GTGACTGAGG CTAGAAGCTG CCATCTTGCC ATGGCCCCGA ATCATGCCGT GGTGTCTCGG 2160
ATGGATAAGG TGGAACGCCT GAAACAGGTG CTGCTCCACC AACAGGCTAA ATTTGGGAGA 2220
AATGGATCTG ACTGCCCGGA CAAGTTTTGC TTATTCCAGT CTGAAACCAA AAACCTTCTG 2280
TTCAATGACA ACACTGAGTG TCTGGCCAGA CTCCATGGCA AAACAACATA TGAAAAATAT 2340
TTGGGACCAC AGTATGTCGC AGGCATTACT AATCTGAAAA AGTGCTCAAC CTCCCCCCTC 2400
CTGGAAGCCT GTGAATTCCT CAGGAAGTAA AACCGAAGAA GATGGCCCAG CTCCCCAAGA 2460
AAGCCTCAGC CATTCACTGC CCCCAGCTCT TCTCCCCAGG TGTGTTGGGG CCTTGGCTCC 2520
CCTGCTGAAG GTGGGGATTG CCCATCCATC TGCTTACAAT TCCCTGCTGT CGTCTTAGCA 2580
AGAAGTAAAA TGAGAAATTT TGTTGATATT CAAAAAAAA 2619

SEQ ID NO: 15
SEQUENCE LENGTH: 892
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

TCTTGACCGG CACACACAGC TCGCTTCTTC ACTTTCTTTT CCATCCACTG CCGGACCCAA 60
GCCAGCCTTC CAGGGAGCAG CCATGCCTTA CCTCTACCGG GCCCCAGGGC CTCAGGCACA 120
CCCGGTTCCC AAGGACGCCC GGATCACCCA CTCCTCAGGC CAGARCTTTG ARCAAATGAA 180
GCAGGARTGC CTGCAGARAR GCACCCTGTT TGAGGATGCA GACTTCCCAG CCAGCAATTC 240
CTCCCTGTTC TACAGTGAGA GGCCGCAGAT CCCCTTTGTG TGGAAACGAC CARGGGAAAT 300
CGTGAAAAAC CCARAATTCA TTCTTGGAGG GGCCACCAGG ACTGATATCT GCCAGGGAGA 360
GCTGGGAGAC TGCTGGCTAT TAGCCGCCAT CGCCTCCCTT ACGCTTAATC AAAAAGCACT 420
GGCCAGAGTC ATCCCCCAGG ACCAAAGCTT TGGCCCTGGT TATGCCGGGA TATTCCATTT 480
CCAGTTCTGG CAGCACAGTG AGTGGCTGGA CGTGGTGATC GATGACCGCC TGCCCACCTT 540
CAGGGACCGC TTGGTTTTCC TCCACTCTGC CGACCACAAC GARTTCTGGA RCGCCTTGCT 600
GGAAAAAGCC TACGCCAAGC TAAATGGGAG CTATGAAGCT CTGAAGGGAG GCAGCGCCAT 660
CGAGGCCATG GAAGACTTCA CTGGGGGTGT GGCAGAGACC TTCCAAACTA AAGAGGCCCC 720
CGAGAACTTC TATGAGATTC TAGAGAAGGC TTTGAAGANA NGCTCCCTGC TGGGCTGCTT 780

CATTGATACC AGAAGTGCTG CAGAATCTGA GGCCCGGACG CCGTTTGGTC TTATTAAGGG 840
TCATGCCTAC AGTGTAACGG GAATTGACCA GGTAAGCTTC CGAGGCCAGA GA 892

SEQ ID NO: 16
SEQUENCE LENGTH: 508
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION:

TGGAGAATGC GAGCCGGGTG TTCCAGGCTC TCAGTACAAA GAACANGGAG TTCATTCATN 60
TCAATATAAA NGAGTTCATC CATTNGACAA TGAACATCTG AGGCTGCNTT GTAGAGATGC 120
AGCCTGCCCA GNTGAATCTG GGNTTCTGGA CCTNGACCTT CAGAANTTCT CTTGGTGTGG 180
AACCATTACG CCCAGGGTTC ACTCCCCTCT CATCGTCCGG CCTTCTCCCT TCATCTTGAT 240
CTGGGAAGAA TGAAATGAAC TCAGCTACAC TCTCTGATTT TGTGCTACTC CTTTGTAAAG 300
TCACTGCCTT AAGGGGGCTG ATGGCGCCAC CTGTGCCTTA CATCCAGGTT CAGGCATCAC 360
TAGCTTTCCC ACACTCTACT TTCCTTATTT CCTTCCATTA AGAATTACTC AGAGTTCTAA 420
CGCACAGAAT CCTGACTTCC ATGTAGCTCC AGTCATTGTG ATCAGACATC CTTTATAAAA 480
CATGTTTTTA TAAATGTGTA TGTGGAAT 508

SEQ ID NO: 17
SEQUENCE LENGTH: 194
SEQUENCE TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION:

Ser Val His Cys Phe Arg Glu Asp Lys Met Lys Phe Thr Ile Val
1 5 10 15
Phe Ala Gly Leu Leu Gly Val Phe Leu Ala Pro Ala Leu Ala Asn
20 25 30
Tyr Asn Ile Asn Val Asn Asp Asp Asn Asn Asn Ala Gly Ser Gly
35 40 45
Gln Gln Ser Val Ser Val Asn Asn Glu His Asn Val Ala Asn Val
50 55 60

```
Asp Asn Asn Asn Gly Trp Asp Ser Trp Asn Ser Ile Trp Asp Tyr
                65                  70                  75
Gly Asn Gly Phe Ala Ala Thr Arg Leu Phe Gln Lys Lys Thr Cys
                80                  85                  90
Ile Val His Lys Met Asn Lys Glu Val Met Pro Ser Ile Gln Ser
                95                 100                 105
Leu Asp Ala Leu Val Lys Glu Lys Lys Leu Gln Gly Lys Gly Pro
               110                 115                 120
Gly Gly Pro Pro Pro Lys Gly Leu Met Tyr Ser Val Asn Pro Asn
               125                 130                 135
Lys Val Asp Asp Leu Ser Lys Phe Gly Lys Asn Ile Ala Asn Met
               140                 145                 150
Cys Arg Gly Ile Pro Thr Tyr Met Ala Glu Glu Met Gln Glu Ala
               155                 160                 165
Ser Leu Phe Phe Tyr Ser Gly Thr Cys Tyr Thr Thr Ser Val Leu
               170                 175                 180
Trp Ile Val Asp Ile Ser Phe Cys Gly Asp Thr Val Glu Asn
               185                 190
```

SEQ ID NO: 18
SEQUENCE LENGTH: 51
SEQUENCE TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Val Asp Asp Lys Arg Lys Ser Ala Leu Trp Lys Glu Arg Thr
  1                 5                  10                  15
Val Ser Thr Arg Val Lys Ser Met Asn Ala Ser Ile Glu Arg Thr
                20                  25                  30
Arg Gly Asn Ile Pro Ser Thr Gly Leu His Thr Cys Ile Tyr Ile
                35                  40                  45
Leu Glu Asn Thr Ala Met
                50
```

SEQ ID NO: 19
SEQUENCE LENGTH: 63
SEQUENCE TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Gly Gln Gln His Gln Tyr Tyr Leu Gly Ala Tyr Arg Asn Ala
 1               5                  10                  15
Glu Phe Gln Ala His Cys Arg Ser Thr Glu Ser Lys Ser Ser Phe
            20                  25                  30
Ser Lys Ile Ser Gln Thr Ile Ser Thr Gly Leu His Pro Phe Tyr
            35                  40                  45
Pro Ser Leu Ala Ile Arg Asp Val Arg Ser Glu Gly Phe Lys Arg
            50                  55                  60
Ser Pro Glu
```

SEQ ID NO: 20
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:

```
TCTTTGCTGG ACTTCTTGGA                20
```

SEQ ID NO: 21
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:

```
CTTTGTTTGG GTTGACTGAG                20
```

SEQ ID NO: 22
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CACCCTCATT ACATCATCAG 20

SEQ ID NO: 23
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
ATTCCTTGTG TCTTCTGGTA 20

SEQ ID NO: 24
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CAGTCCTACT TCTCCTATCT C 21

SEQ ID NO: 25
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:

ATCATAGCTC AGACCATACC T 21

SEQ ID NO: 26
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
GATCCTGCAG GACTACAAAT C 21

SEQ ID NO: 27
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
GCCTATATAG AAAAATGAAG 20

SEQ ID NO: 28
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CACCTAGTGA CCGTTCCAGA T 21

SEQ ID NO: 29
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTCATCTCCT TGGGTGTTAT T 21

SEQ ID NO: 30
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CTCAGACGCT CAGGAAATAG A 21

SEQ ID NO: 31
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
AATGGGGGAA GTATGTAGGA G 21

SEQ ID NO: 32
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTACGGATCA TTTCTCTACT C 21

SEQ ID NO: 33
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
AGGGCAAGAT GAAGTGAAAG G 21

SEQ ID NO: 34
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TCCGGAAAGA AGAGCGAGAG A 21

SEQ ID NO: 35
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TGAAACACAA CTACCCCAAT G 21

SEQ ID NO: 36
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
ATAGCAAAGG TAAACTCTCA 20

SEQ ID NO: 37

SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TCAATCAGTA GTTCCCAGTA 20

SEQ ID NO: 38
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTAACAGCCC AATATCTACA 20

SEQ ID NO: 39
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
GAACAAGTGA TTATGCTACC 20

SEQ ID NO: 40
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
AGAATAAGCA ACTTGGAAAA 20

SEQ ID NO: 41
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TGAATCTGAT GACTATGTGC 20

SEQ ID NO: 42
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TCCTGGATAC CTTTTGGACC 20

SEQ ID NO: 43
SEQUENCE LENGTH: 19
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CATCAGGGCT ACAAGGAAA 19

SEQ ID NO: 44
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:

CAGATCTACC GAATCAAAAT C 21

SEQ ID NO: 45
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
ACCAGAATTA GGAATAAGGA T 21

SEQ ID NO: 46
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
GACTCCATGG CAAAACAACA 20

SEQ ID NO: 47
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TCTTCTTCGG TTTTACTTCC 20

SEQ ID NO: 48
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
AGGCACCAGG GCGTGATGGT 20

SEQ ID NO: 49
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
GGTCTCAAAC ATGATCTGGG 20

SEQ ID NO: 50
SEQUENCE LENGTH: 10
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CTTGATTGCC 10

SEQ ID NO: 51
SEQUENCE LENGTH: 10
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
AGGTGACCGT 10

SEQ ID NO: 52
SEQUENCE LENGTH: 10
SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
GTTGCGATCC 10

SEQ ID NO: 53
SEQUENCE LENGTH: 10
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CTGATCCATG 10

SEQ ID NO: 54
SEQUENCE LENGTH: 10
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
CTGCTTGATG 10

SEQ ID NO: 55
SEQUENCE LENGTH: 10
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
GATCTGACTG 10

SEQ ID NO: 56

SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TAA 13

SEQ ID NO: 57
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TAC 13

SEQ ID NO: 58
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TAG 13

SEQ ID NO: 59
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TCA 13

SEQ ID NO: 60
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TCC 13

SEQ ID NO: 61
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TCG 13

SEQ ID NO: 62
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TGA 13

SEQ ID NO: 63
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:

TTTTTTTTTT TGC 13

SEQ ID NO: 64
SEQUENCE LENGTH: 13
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid synthetic DNA
SEQUENCE DESCRIPTION:
TTTTTTTTTT TGG 13

SEQ ID NO: 65
SEQUENCE LENGTH: 264
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
SEQUENCE DESCRIPTION:
AGGCACCAGG GCGTGATGGT GGGCATGGGT CAGAAGGATT CCTATGTGGG CGACGAGGCC 60
CAGAGCAAGA GAGGCATCCT CACCCTGAAG TACCCCATCG AGCACGGCAT CGTCACCAAC 120
TGGGACGACA TGGAGAAAAT CTGGCACCAC ACCTTCTACA ATGAGCTGCG TGTGGCTCCC 180
GAGGAGCACC CCGTGCTGCT GACCGAGGCC CCCCTGAACC CCAAGGCCAA CCGCGAGAAG 240
ATGACCCAGA TCATGTTTGA GACC 264

Claims

1. A method for detecting a cancer cell in a resected specimen characterized by determining a change in an expression level of at least one of cancer-associated genes selected from genes of which cDNA is a DNA comprising a nucleotide sequence as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing, or a DNA capable of hybridizing with a nucleic acid consisting of a nucleotide sequence as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing under stringent conditions.

2. The detection method according to claim 1, characterized in that the change in an expression level of a cancer-associated gene is determined by the change in an expression level of mRNA corresponding to said gene.

3. The detection method according to claim 2, characterized in that the change in an expression level of mRNA is detected by utilizing a nucleic acid amplification method based on said mRNA or a partial portion thereof.

4. The detection method according to claim 3, characterized in that said nucleic acid amplification method is polymerase chain reaction.

5. The detection method according to claim 2, characterized in that the change in an expression level of mRNA is detected by Northern hybridization method.

6. The detection method according to claim 2, characterized in that the change in an expression level of mRNA is

detected by RNase protection assay.

7. The detection method according to claim 1, characterized in that the change in an expression level of a cancer-associated gene is determined by a change in an expression level of a protein encoded by said gene.

8. The detection method according to claim 7, characterized in that the change in expression of the protein is detected by utilizing an antibody capable of recognizing said protein.

9. A kit for detecting cancer by the method of claim 3, wherein the kit comprises primers for amplifying mRNA of which change in an expression level is to be determined or a partial portion thereof.

10. A kit for detecting a cancer cell by the method of claim 5 or 6, wherein the kit comprises a probe capable of hybridizing with mRNA of which change in an expression level is to be determined.

11. A kit for detecting a cancer cell by the method of claim 8, wherein the kit comprises an antibody recognizing a protein of which change in an expression level is to be determined.

12. A method for controlling proliferation of a cancer cell using a substance specifically binding to a gene or an expression product of said gene, characterized in that cDNA of the gene is at least one of DNAs selected from a DNA comprising a nucleotide sequence as shown in any one of SEQ ID NOs: 1 to 16 in Sequence Listing, or a DNA capable of hybridizing with a nucleic acid consisting of any one of these nucleotide sequences under stringent conditions, wherein the DNA is usable for detection of a cancer cell by a change in an expression level thereof.

13. A peptide usable for detection of a cancer cell, characterized in that the peptide is shown in an amino acid sequence comprising an entire portion of an amino acid sequence as shown in SEQ ID NOs: 17 to 19 in Sequence Listing or a partial portion thereof.

14. A peptide usable for detection of a cancer cell, characterized in that the peptide has an amino acid sequence comprising an amino acid sequence resulting from at least one of deletion, substitution or addition of one or more amino acid residues in the amino acid sequence as shown in SEQ ID NOs: 17 to 19 in Sequence Listing.

15. A nucleic acid encoding the peptide of claim 13 or 14.

16. The nucleic acid according to claim 15, characterized in that the nucleic acid has a nucleotide sequence comprising the nucleotide sequence as shown in any one of sequences of SEQ ID NOs: 1, 2, 9, 11, 12, 13, 15, or 16 in Sequence Listing.

17. A nucleic acid usable for detection of a cancer cell by a change in an expression level thereof, wherein the nucleic acid is capable of hybridizing with the nucleic acid of claim 15 or 16 under stringent conditions.

18. An antibody usable for detection of a cancer cell, wherein the antibody recognizes the peptide of claim 13 or 14.

FIG. 1

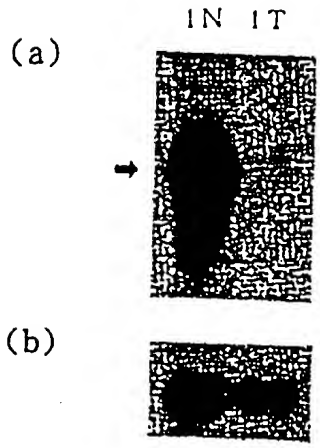

FIG. 2

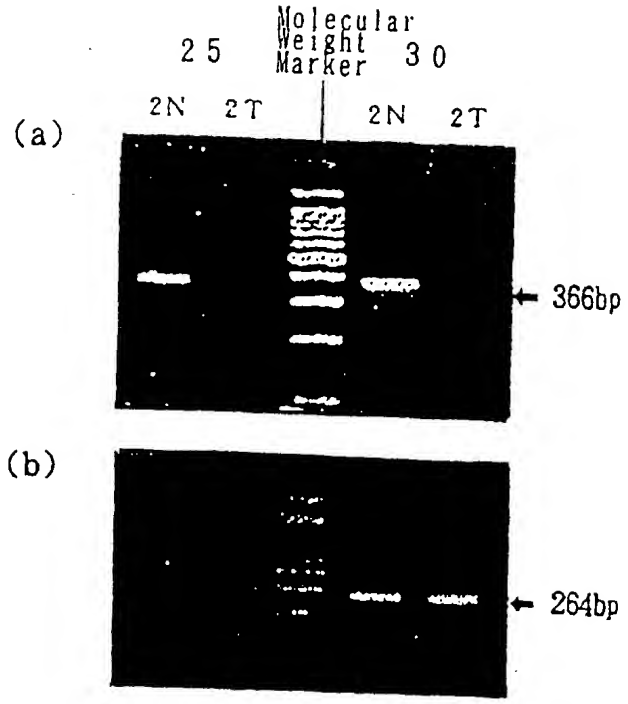

FIG. 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP98/00667

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[6] C12N15/11, C12Q1/68, C07K14/47, C07K16/18 //
(C12N15/11, C12R1:91)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[6] C12N15/11, C12Q1/68, C07K14/47, C07K16/18

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
BIOSIS (DIALOG), WPI (DIALOG), GenBank/EMBL/DDBJ (GENETYX)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SALESIOTIS, A.N., et al., "Identification of novel genes from stomach cancer cell lines by differential display", Cancer Letters (1995) Vol. 91, No. 1, p.47-54	1-11, 13-18
A	WANG, F.L., et al., "Two differentially expressed genes in normal human prostate tissue and in carcinoma", Cancer Research (1996) Vol. 56, No. 16, p.3634-3637	1-11, 13-18
A	WATSON, M.A., et al., "Isolation of differentially expressed sequence tags from human breast cancer", Cancer Research (1994) Vol. 56, No. 17, p.4598-4602	1-11, 13-18

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
May 21, 1998 (21. 05. 98)	June 2, 1998 (02. 06. 98)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)